The 21st Century
Earth Science COE Program
University of Tokyo

GeoFEM

Integrated Predictive Simulation System for Earthquake and Tsunami Disaster

Current and Future Contributions from the HPC Community

Kengo Nakajima

The 21st Century Earth Science COE Program
Department of Earth & Planetary Science
The University of Tokyo and CREST/JST

2007 Workshop on Community Finite Element Models for Fault Systems and Tectonic Studies, Colorado School of Mines, Golden, Colorado, June 27, 2007.

Acknowledgements

- Computational Infrastructure for Geodynamics (CIG)
- The 21st Century Earth Science COE (Center of Excellence) Program
  - Department of Earth & Planetary Science, The University of Tokyo
- CREST, Japan Science and Technology Agency (JST)
- Colleagues
  - GeoFEM, HPC-MW, CREST etc. Projects
  - Dr. Mamoru Hyodo (ESC, Japan)

Overview of this talk

- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- Overview of the Current Project by JST
  - Integrated Predictive Simulation System for Earthquake and Tsunami Disaster
- Some Technical Issues
  - Parallel Preconditioning Methods
  - Vector vs. Scalar Processors
  - Parallel Programming Models in Multi-Core Era
- Future Directions

GeoFEM: FY.1998-2002

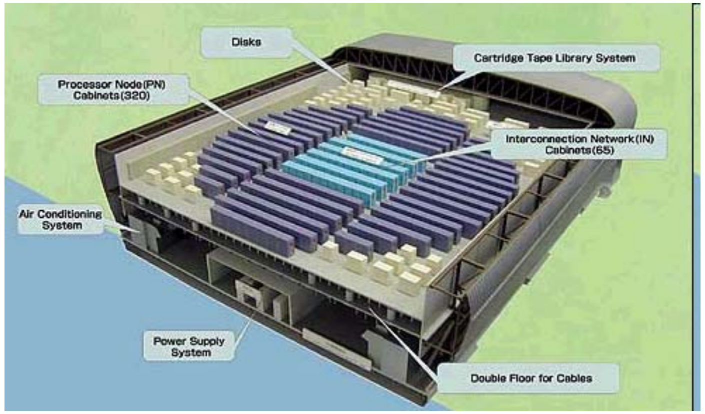
http://geofem.tokyo.rist.or.jp/

- Parallel FEM **platform (or framework)** for solid earth simulation.
  - parallel I/O, **parallel linear solvers**, parallel visualization
  - solid earth: earthquake, plate deformation, mantle/core convection, etc.
- Part of national project by STA/MEXT for large-scale earth science simulations using the Earth Simulator.
- Strong collaborations between HPC and natural science (solid earth) communities.
- Current Activity
  - Supporting several users in ESC (Earth Simulator Center)

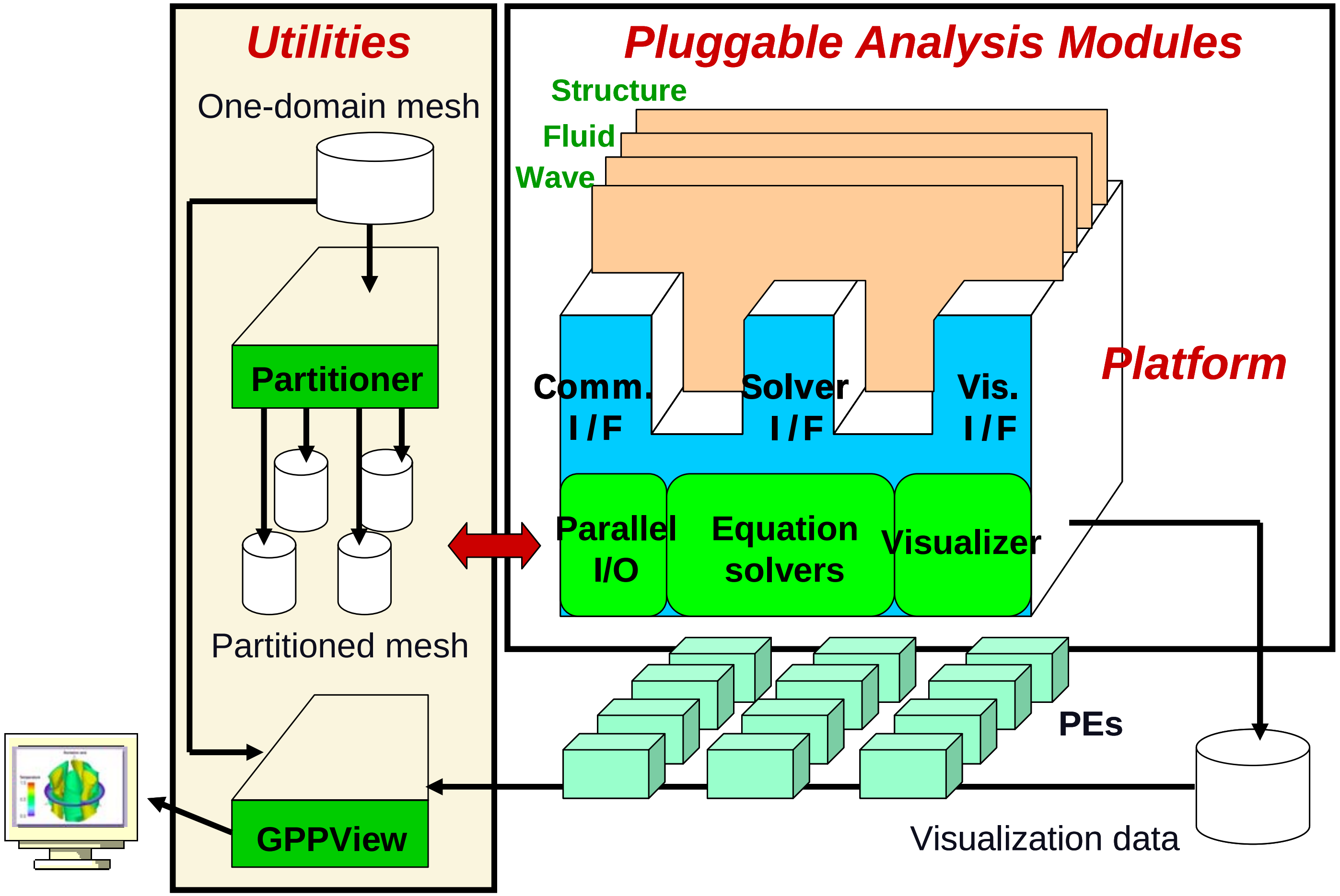
Earth Simulator (ES)

http://www.es.jamstec.go.jp/

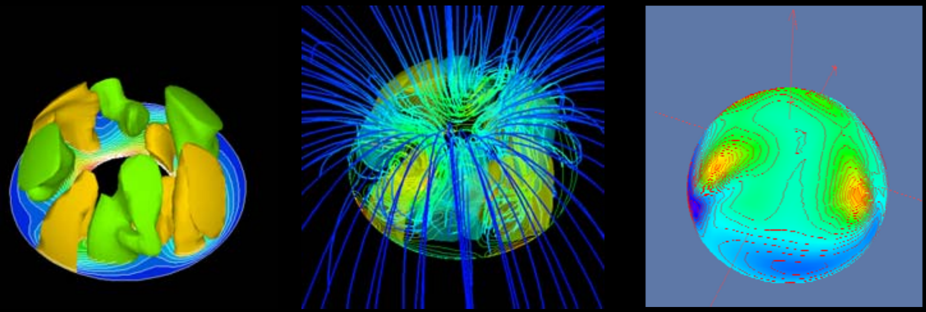
- 640×8= 5,120 Vector Processors
  - **SMP Cluster-Type Architecture**
  - 8 GFLOPS/PE
  - 64 GFLOPS/Node
  - 40 TFLOPS/ES
- 16 GB Memory/Node, 10 TB/ES
- 640×640 Crossbar Network
  - 12.3 GB/sec×2
- Memory BWTH with 32 GB/sec.
- **35.6 TFLOPS for LINPACK (2002-March)**
  - **14th in Nov.06 list (Jun.07 list this week)**
- **26 TFLOPS for AFES (Climate Simulation)**



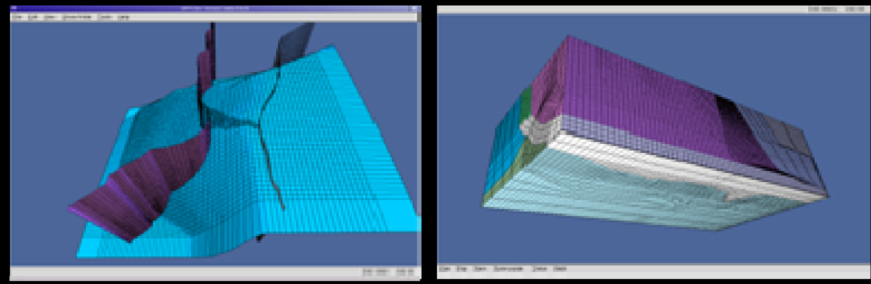
System Configuration of GeoFEM



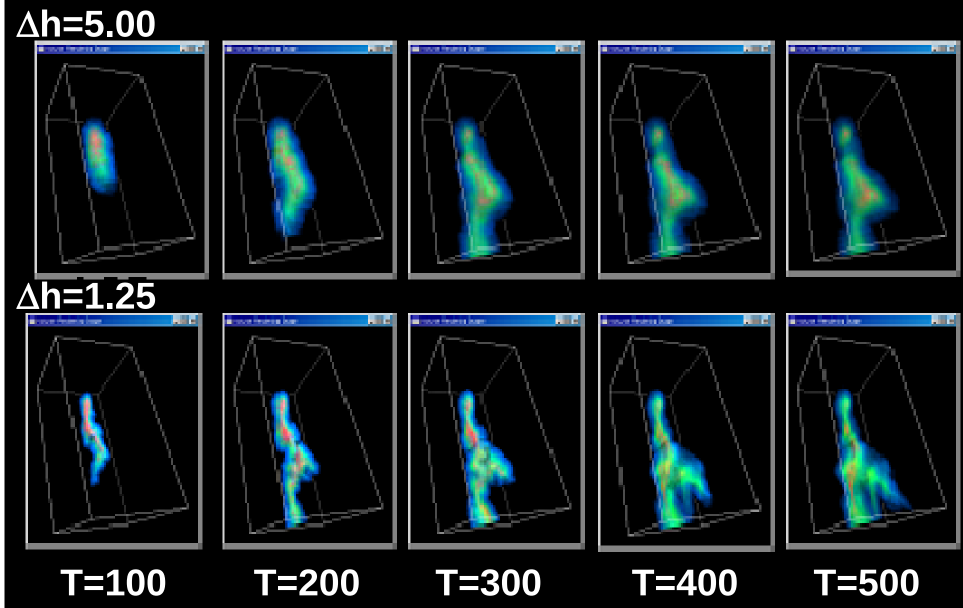
Results on Solid Earth Simulation



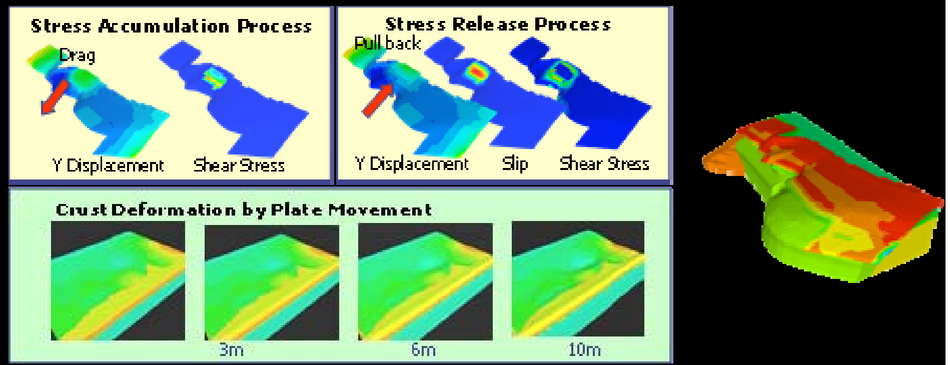
Magnetic Field of the Earth : MHD code



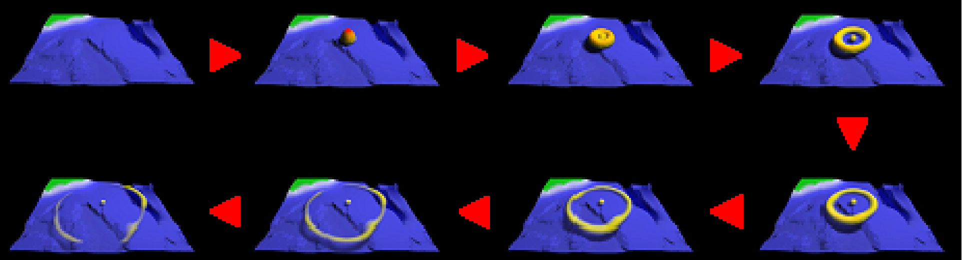
Complicated Plate Model around Japan Islands



Transportation by Groundwater Flow through Heterogeneous Porous Media

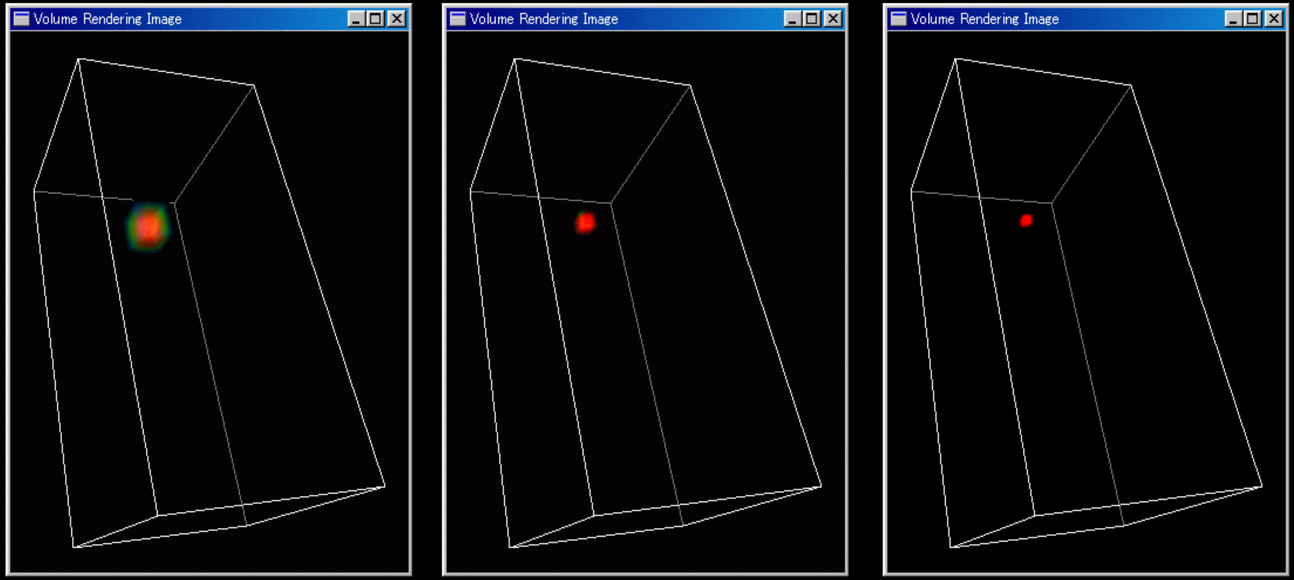


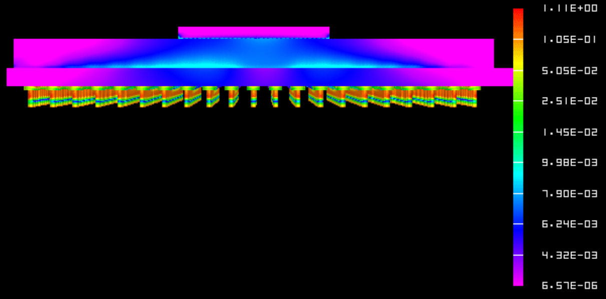
Simulation of Earthquake Generation Cycle in Southwestern Japan

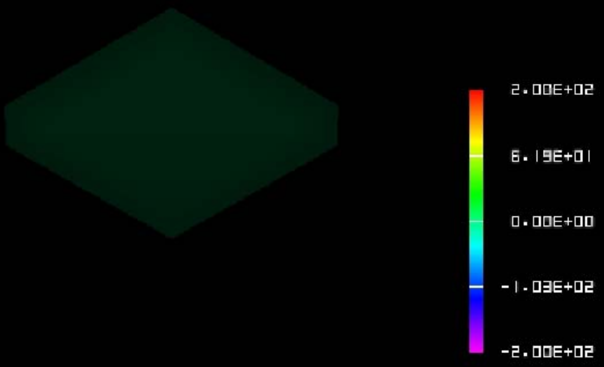


TSUNAMI !!

Results by GeoFEM



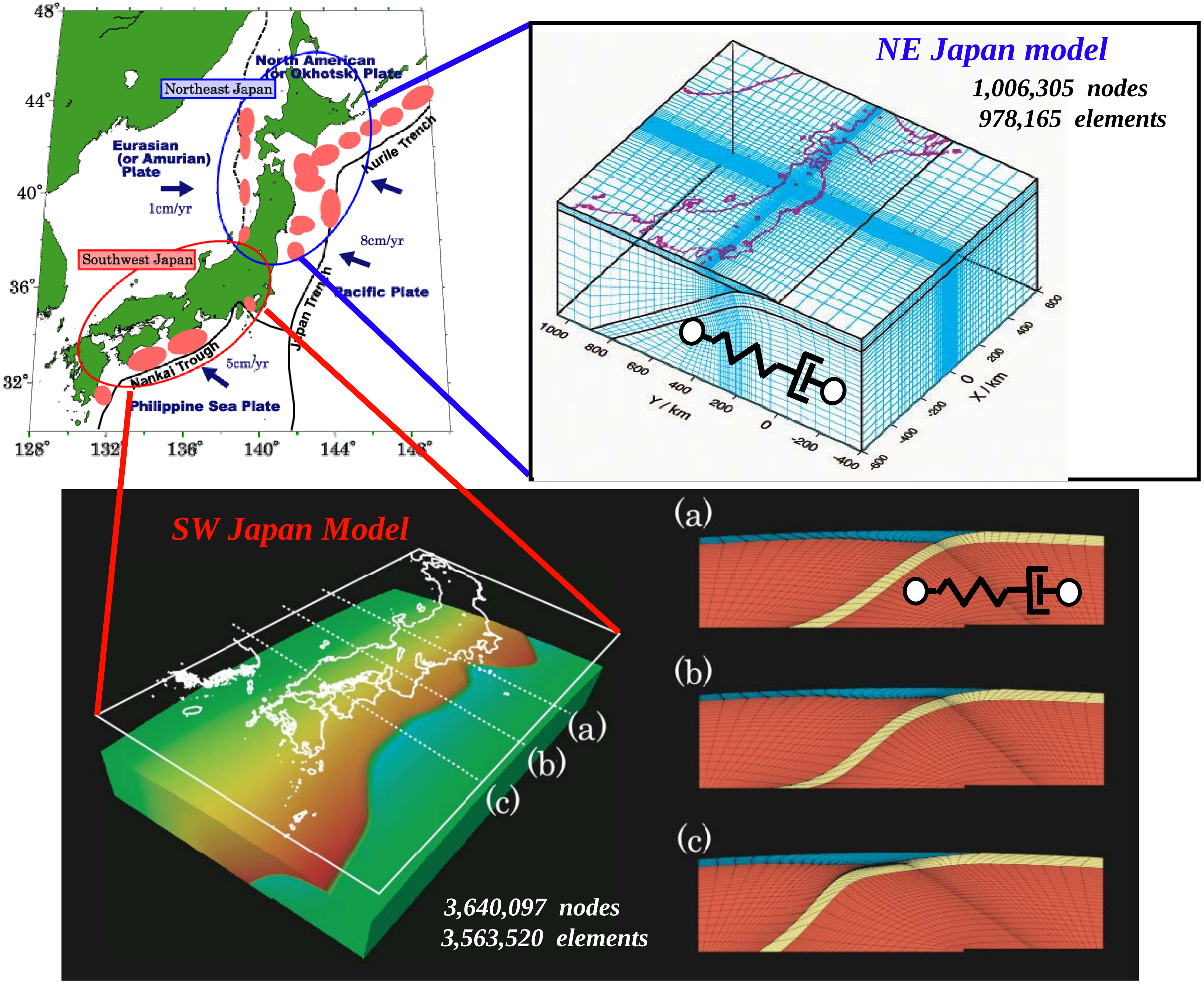




Quasi-static crustal deformation due to an internal dislocation source in multilayered elastic/viscoelastic zones using GeoFEM framework

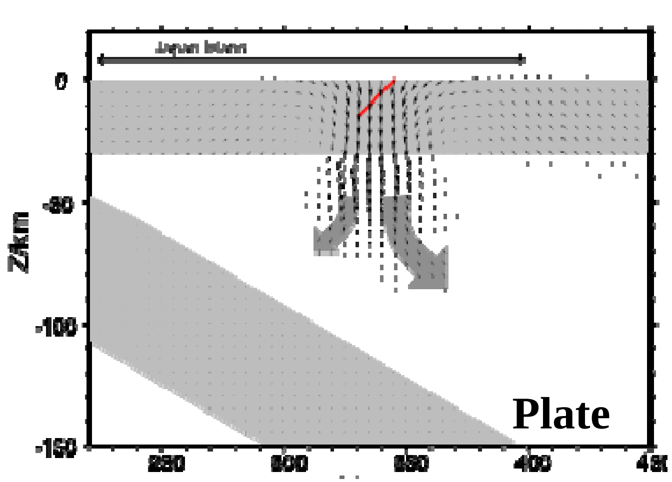
- Dr. Mamoru Hyodo (Earth Simulator Center, JAMSTEC)

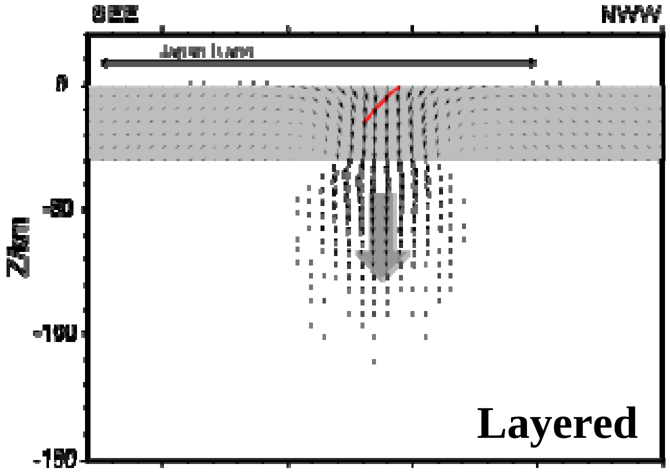
Regional Viscoelastic FE models in Japan



Simulation of PostSeismic Deformation in NE Japan

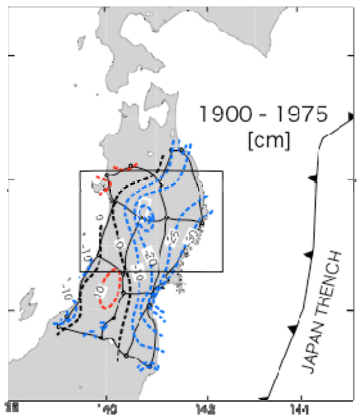
Q. Subducting plate (configuration) controls postseismic deformation ?

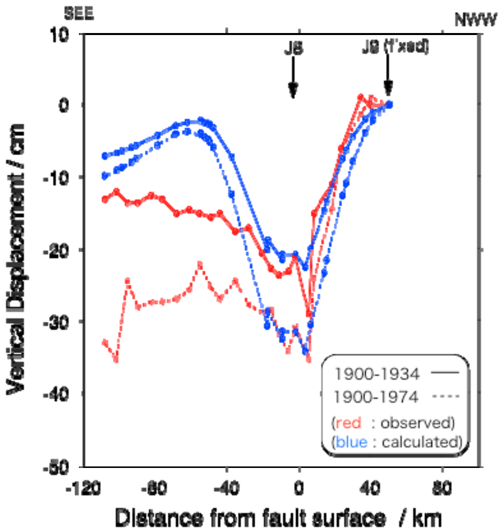


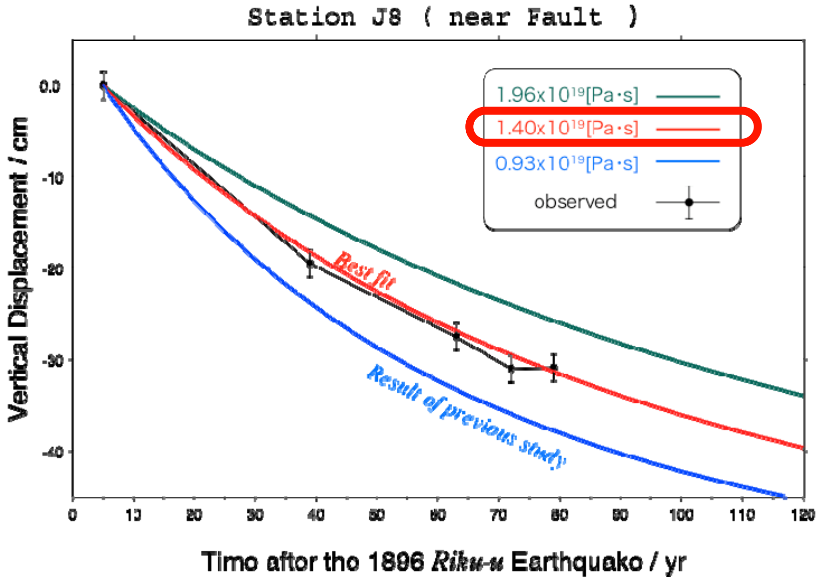


A. No. Not so dominant.

Q. How viscid is the asthenosphere beneath NE Japan ?

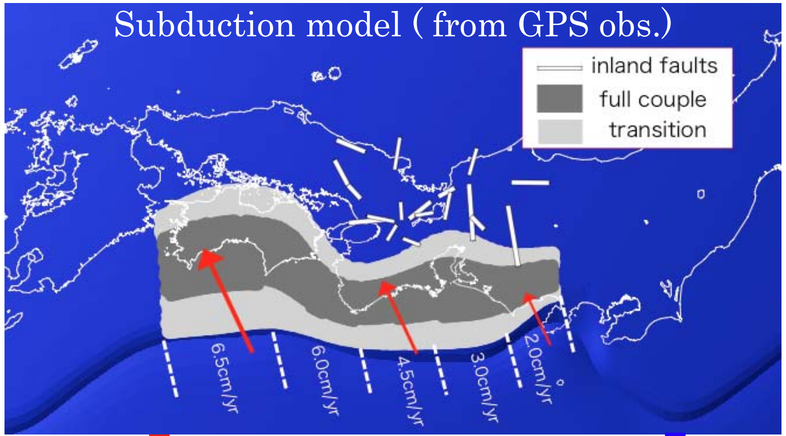


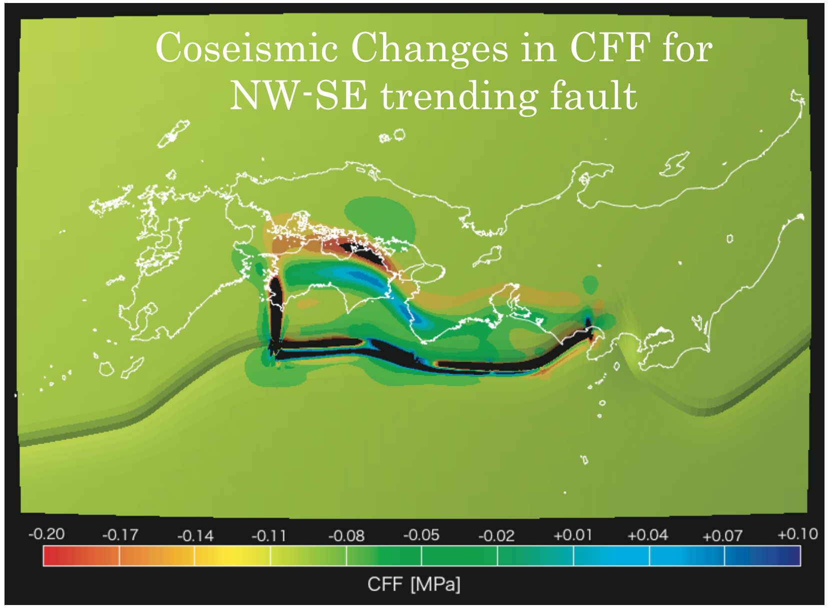




Kinematic Simulation of Earthquake cycles in SW Japan

Q. How great earthquakes along the Nankai trough affect stress state on inland faults ?

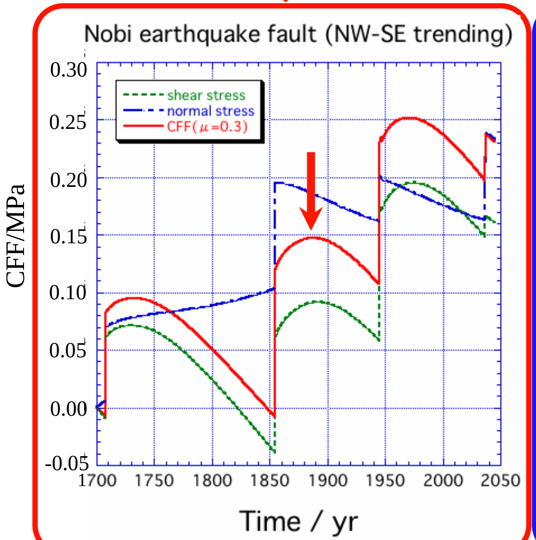


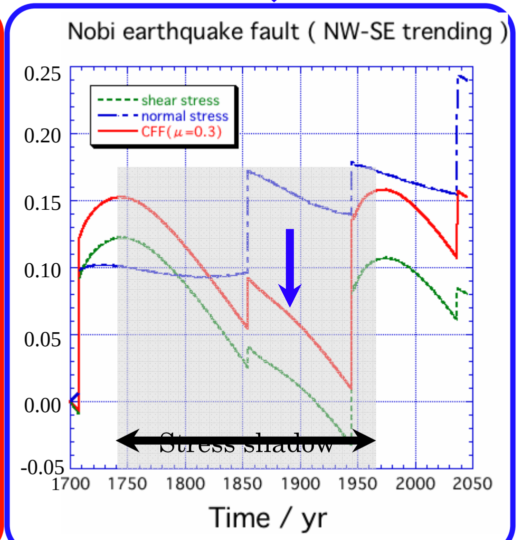


+ Slip predictable earthquake occurrence

+ Time predictable earthquake occurrence

Which is better?





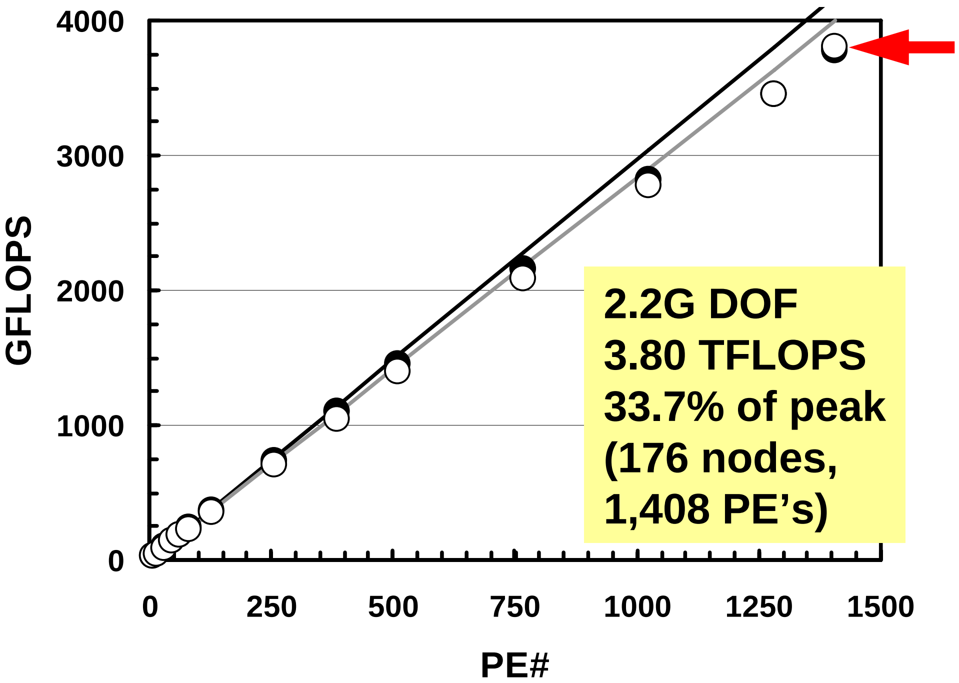
- ***Slip predictable model*** ***can successfully explains the occurrence of many historical inland earthquakes.***
- ***But*** ***time predictable model*** ***cannot...***

Performance of GeoFEM's ICCG Solver Weak Scaling

● Flat-MPI DJDS
○ Hybrid DJDS
— Flat-MPI(ideal)
— Hybrid (ideal)

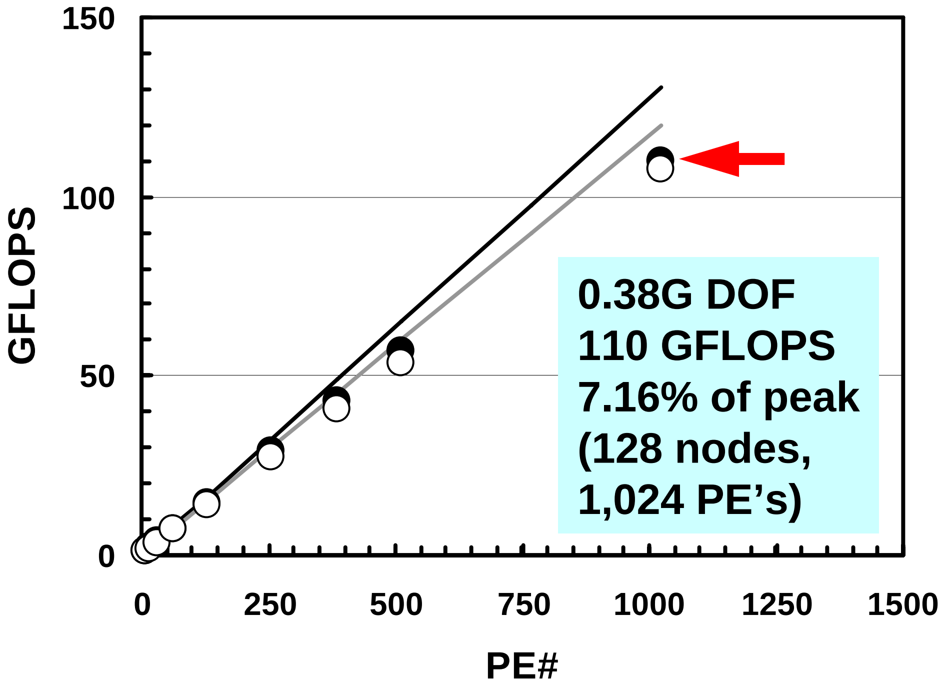
Earth Simulator

1,572,864 DOF/PE
(=3x128x64x64)



IBM SP-3 (Seaborg at LBNL)

375,000 DOF/PE
(=3x50^3)



The 21st Century Earth Science COE Program, EPS Dept., University of Tokyo

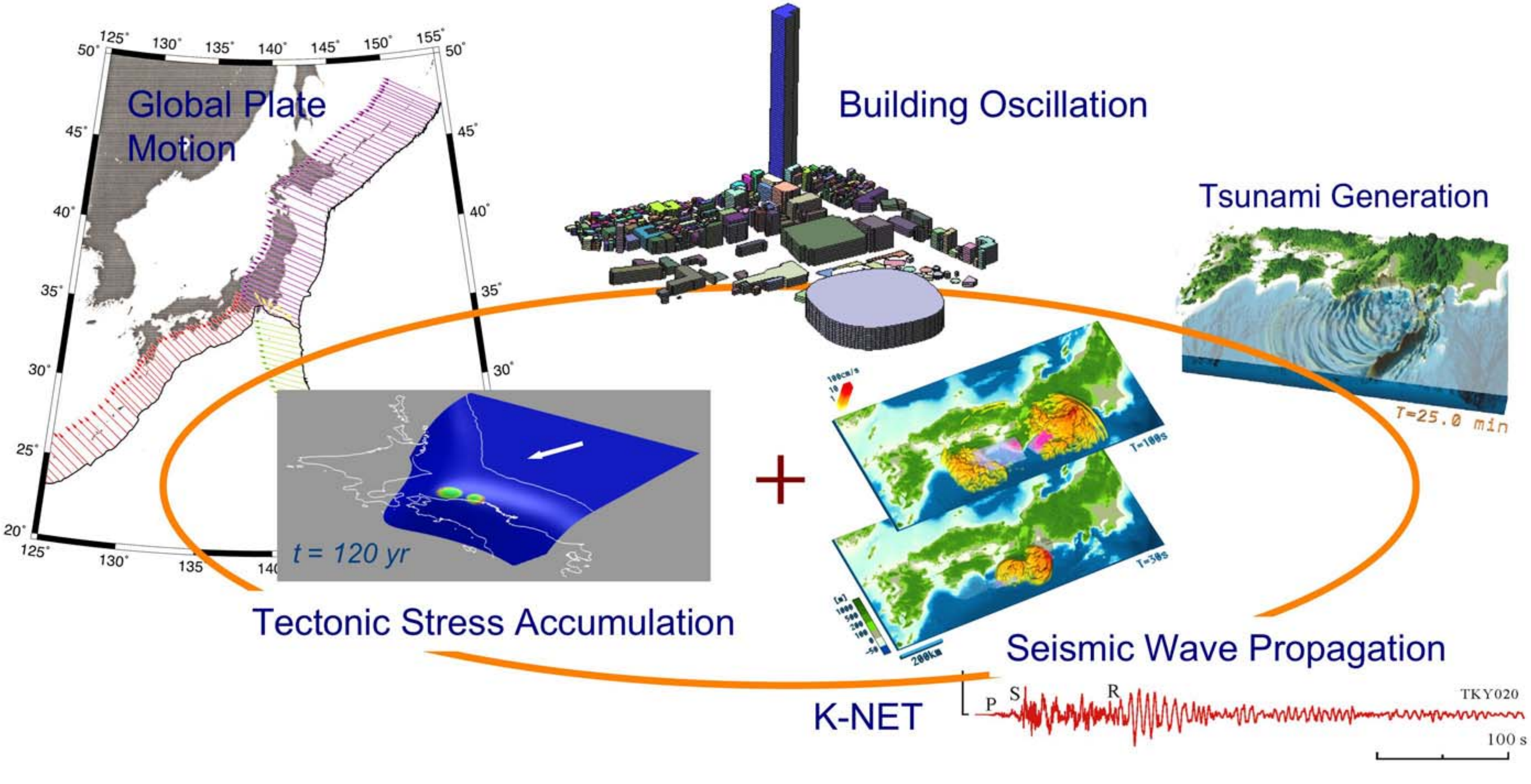
Oct.2003-Mar.2008

- http://www.eps.s.u-tokyo.ac.jp/jp/COE21/eng/
- COE = **C**enter **O**f **E**xcellence
  - One of 6 COE's for Earth Science
- Predictability of the Evolution and Variation of the Multi-scale Earth System
  - An Integrated COE for Observational and Computational Earth Science
  - to build an advanced research and education system that can promote study of future variations predictability of a multi-sphere earth system effectively
- I am teaching HPC and Parallel Programming.
  - Only classes on programming of scientific computing in Japan

- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- **Overview of the Current Project by JST**
  - **Integrated Predictive Simulation System for Earthquake and Tsunami Disaster**
- Some Technical Issues
  - Parallel Preconditioning Methods
  - Vector vs. Scalar Processors
  - Parallel Programming Models in Multi-Core Era
- Future Directions

Integrated Predictive Simulation System for Earthquake and Tsunami Disaster

Our goal is to develop an integrated simulation system for predicting earthquake and tsunami disaster, which covers the earthquake-related multi-scale processes from plate motion to building oscillation.



Integrated Predictive Simulation System for Earthquake and Tsunami Disaster

- PI (Principal Investigator)
  - Prof. Mitsuhiro Matsu'ura (The University of Tokyo)
- 5-year National Project
  - Oct. 2005 - Mar. 2011
  - CREST, Japan Science and Technology Agency (JST)
    - Basic Research Programs
    - Category: Integrated Simulations for Multiscale/Multiphysics

Institutions

The 21st Century Earth Science COE Program University of Tokyo

Research into Artifacts, Center for Engineering The University of Tokyo

- The University of Tokyo
  - Department of Earth & Planetary Science (EPS)
  - Earthquake Research Institute (ERI)
  - Research into Artifacts, Center for Engineering (RACE)
- Tokyo Institute of Technology (Titech)
- Sophia University
- National Research Institute for Earth Science and Disaster Prevention (NIED)
- Geographical Survey Institute, Japan (GSI)

東京工業大学 Tokyo Institute of Technology

上智大学 SOPHIA UNIVERSITY

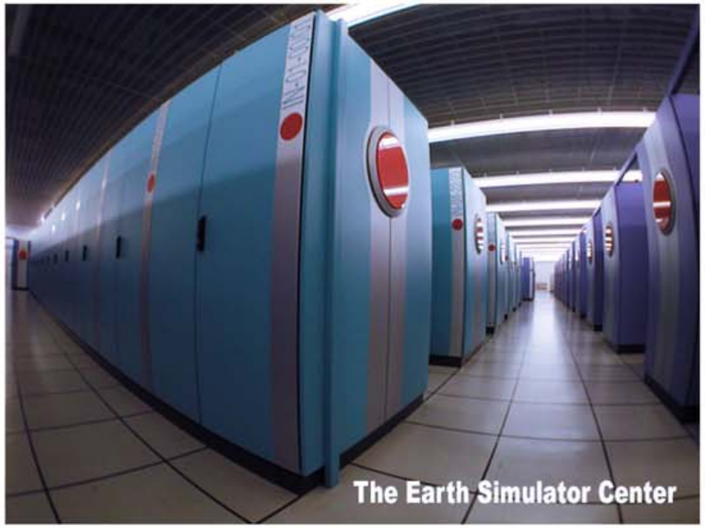
National Research Institute for Earth Science and Disaster Prevention. NIED

国土地理院

Scientific Background

- Solid Earth Simulator Project (1998-2003)
  - Development of the basic simulation models for plate motion, tectonic stress accumulation, earthquake rupture propagation, seismic wave propagation, and building oscillation
  - Development of the solid Earth simulation platform “GeoFEM”

- Solid Earth Simulation Consortium /
  Earth Simulator Research Project (2002-)



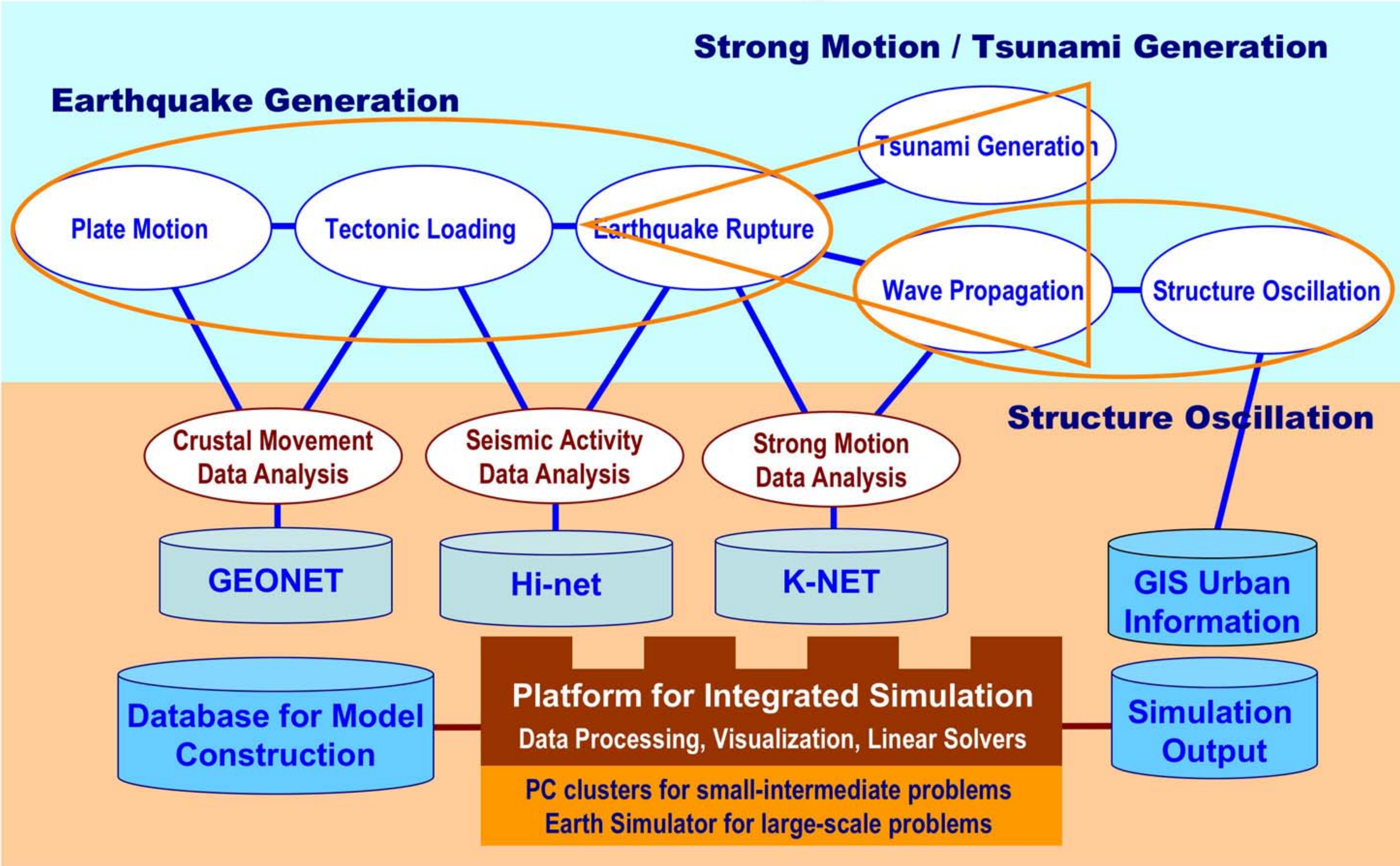
APEC Cooperation for Earthquake Simulation
July 9-14, 2004, Beijing China

- APEC Cooperation for Earthquake Simulation (1998-)

Features of the Project (1/3)

- The first integrated simulation system for prediction of earthquake and tsunami disasters, which covers entire series of multi-scale processes, such as:
  - Plate deformation
  - Dynamic fault rupture
  - Seismic wave/tsunami propagation
  - Oscillation of buildings
- Target hardware
  - Earth Simulator (ES), PC clusters

Outlines of the Integrated Predictive Simulation System

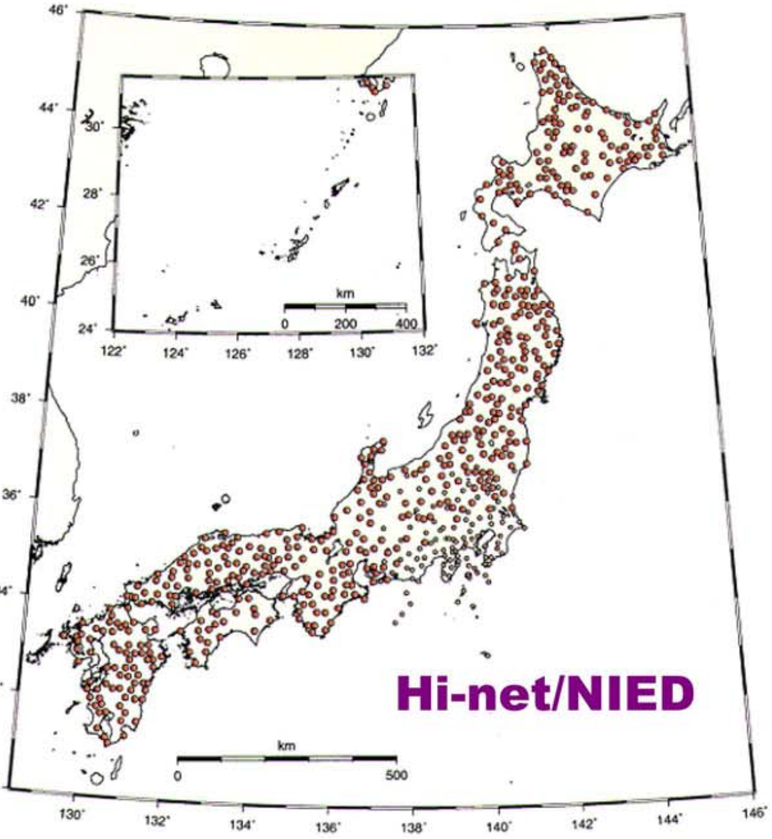


Features of the Project (2/3)

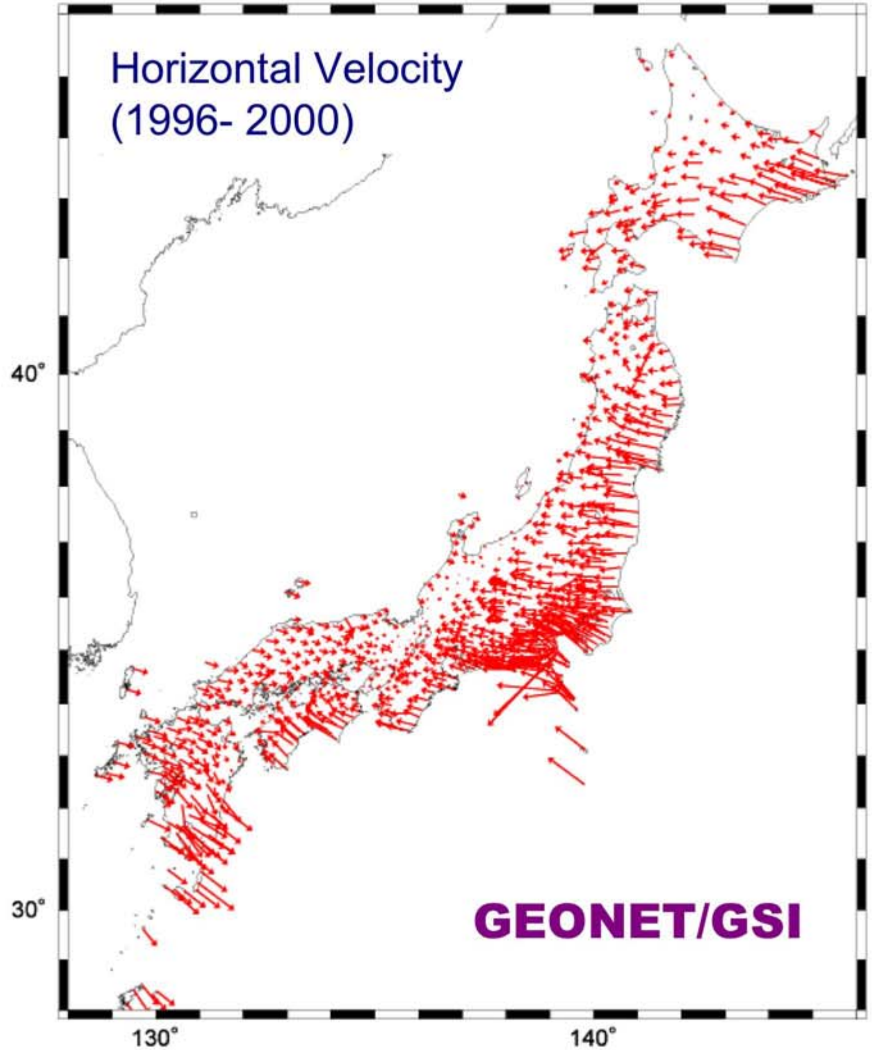
- The simulation is complimented by large number of observational data sets obtained through nation-wide network of seismic instruments and GPS etc.
  - Crustal movement: GEONET
  - Seismic activity: Hi-net
  - Strong Motion: K-NET/KiK-net

- >1,000 points for each network
  - every 20 km

Nation-Wide Seismic and Geodetic Networks In Japan

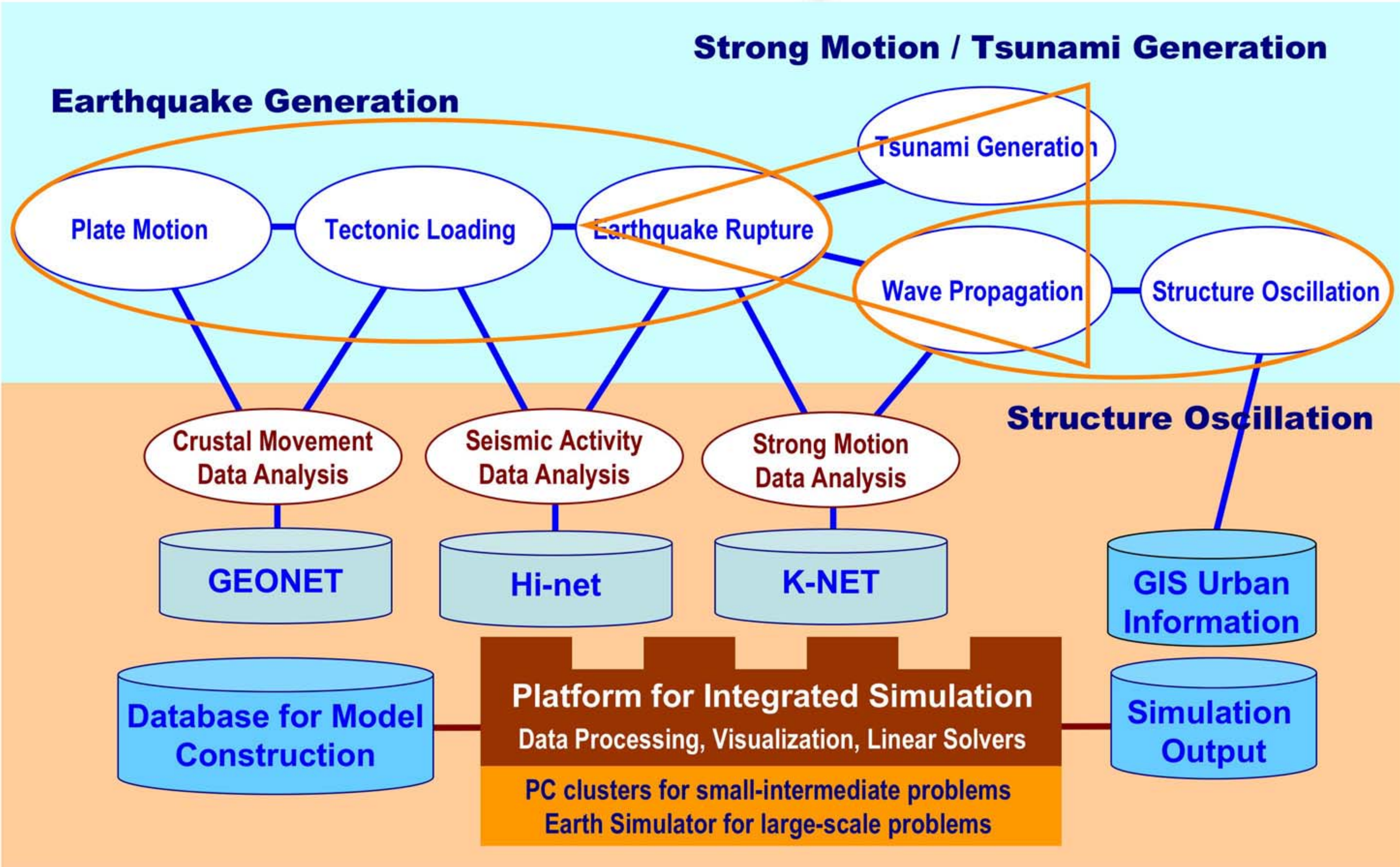


NIED: National Research Institute for Earth Science and Disaster Prevention



GSI: Geographical Survey Institute

Outlines of the Integrated Predictive Simulation System



Features of the Project (3/3)

- Both of advanced simulation models and infrastructure for large-scale simulations are developed: GeoFEM's experience
  - Platform
  - GRID-like Environment
- Strong Collaboration/Integration between ...
  - Simulation and Observation
  - Geophysics and Computer/Computational Science
  - Earthquake Science and Earthquake Engineering
  - Simulation models
  - Observation/Data assimilation
  - Infrastructure (Platform+Database)

Challenges

- Integration of simulation and observation
  - Inversion of geodetic and seismic data
- Coupled simulations

Predictive Simulation for Earthquake Generation at Plate Interfaces

- *At the present stage, given the past fault slip history and the present stress state, we can predict the next-step seismic or aseismic fault-slip motion at plate interfaces through computer simulation.*

- *The inversion analysis of geodetic data and seismic data gives us most reliable information about the past fault slip history and the present stress state at plate interfaces.*

Matsu'ura (2007)

Integration of
simulation & observation

Geodetic Data Inversion to Estimate Slip Perturbation at Plate Interfaces

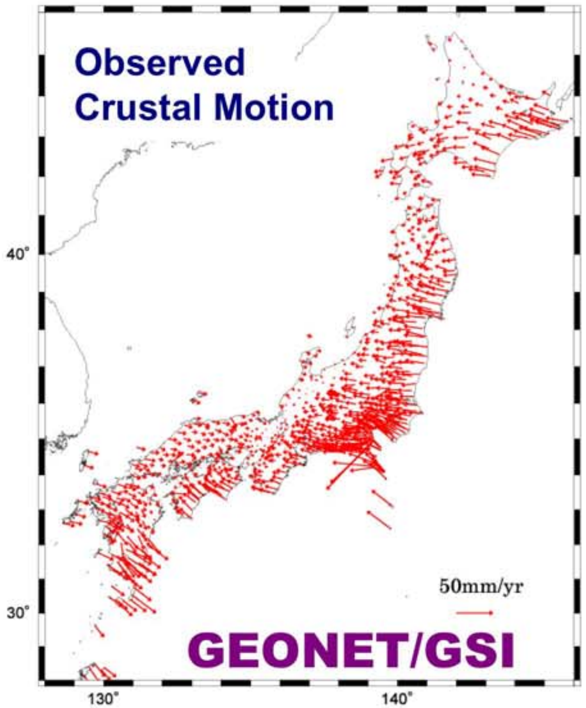
$w(\mathbf{x},t) = V_{pl}t + w_s(\mathbf{x},t)$ Fault slip at a plate interface = steady slip + its perturbation

$$u_i(\mathbf{x},t) = \underbrace{V_{pl}t\int_{\Sigma} U_i(\mathbf{x},\infty;\boldsymbol{\xi},0)\, d\Sigma}_{\text{Crustal deformation due to steady plate subduction}} + \underbrace{\int_0^t d\tau \int_{\Sigma} U_i(\mathbf{x},t-\tau;\boldsymbol{\xi},0)\frac{\partial}{\partial\tau}w_s(\boldsymbol{\xi},\tau)\, d\Sigma}_{\text{Crustal deformation due to slip perturbation}}$$

Crustal deformation due to steady plate subduction

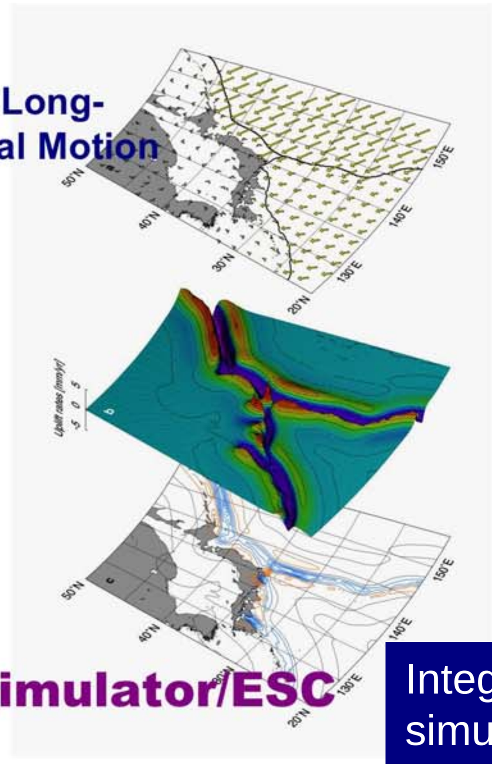
Crustal deformation due to slip perturbation

Observed Crustal Motion



GEONET/GSI

− Computed Long-term Crustal Motion



Earth Simulator/ESC

= Residual Crustal Motion

↓ **Inversion**

Slip Perturbation

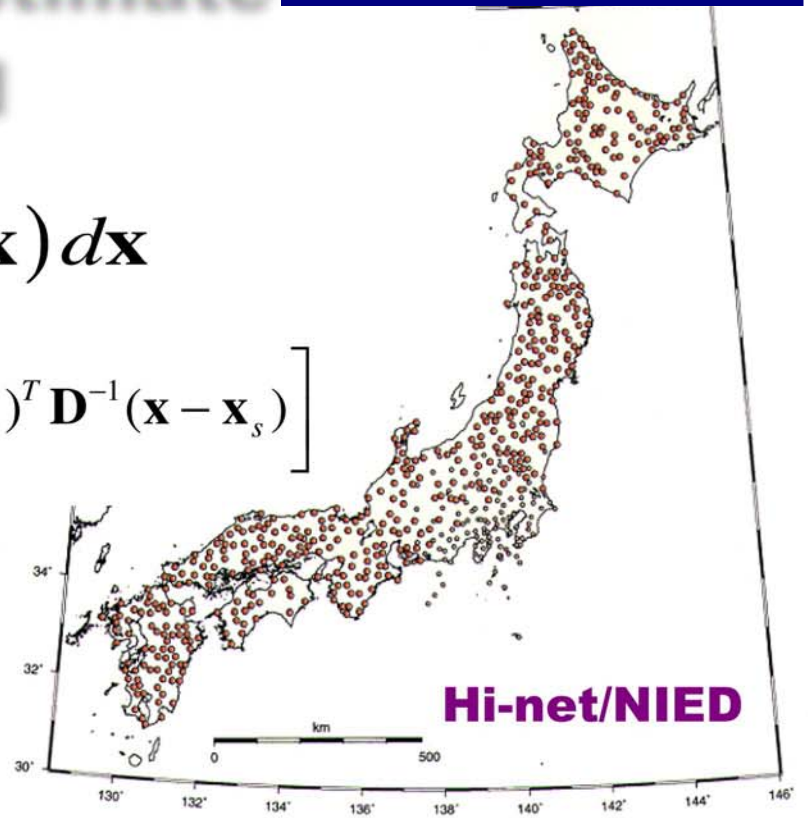
Integration of simulation & observation

Integration of simulation & observation

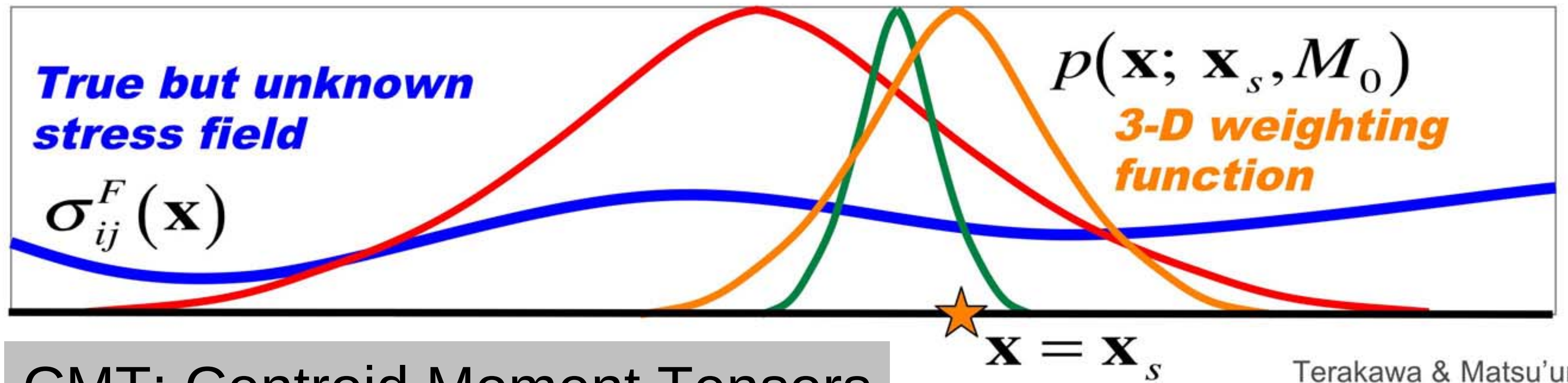
CMT Data Inversion to Estimate Seismogenic Stress Field

$$M_{ij}(\mathbf{x}_s) = \int_V p(\mathbf{x};\ \mathbf{x}_s, M_0)\sigma_{ij}^F(\mathbf{x})\,d\mathbf{x}$$

$$p(\mathbf{x};\ \mathbf{x}_s, M_0) = \frac{M_0}{(2\pi L^2)^{3/2}}|\mathbf{D}|^{-1/2}\exp\left[-\frac{1}{2L^2}(\mathbf{x}-\mathbf{x}_s)^T\mathbf{D}^{-1}(\mathbf{x}-\mathbf{x}_s)\right]$$



Earthquake occurrence can be regarded as a stress release process in the earth's crust.



CMT: Centroid Moment Tensors

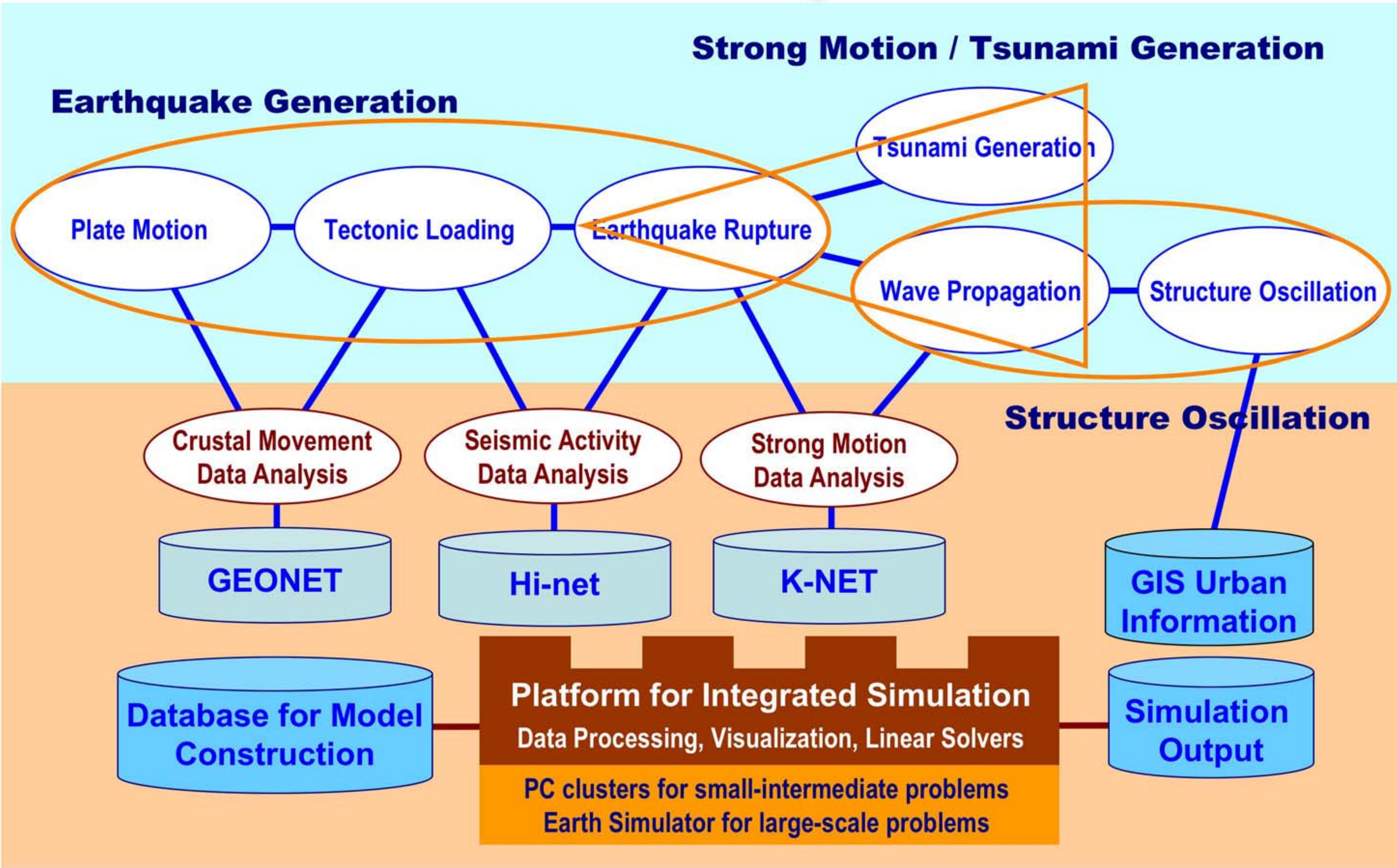
Terakawa & Matsu'ura (2006)

Challenges

- Integration of simulation and observation
  - Inversion of geodetic and seismic data

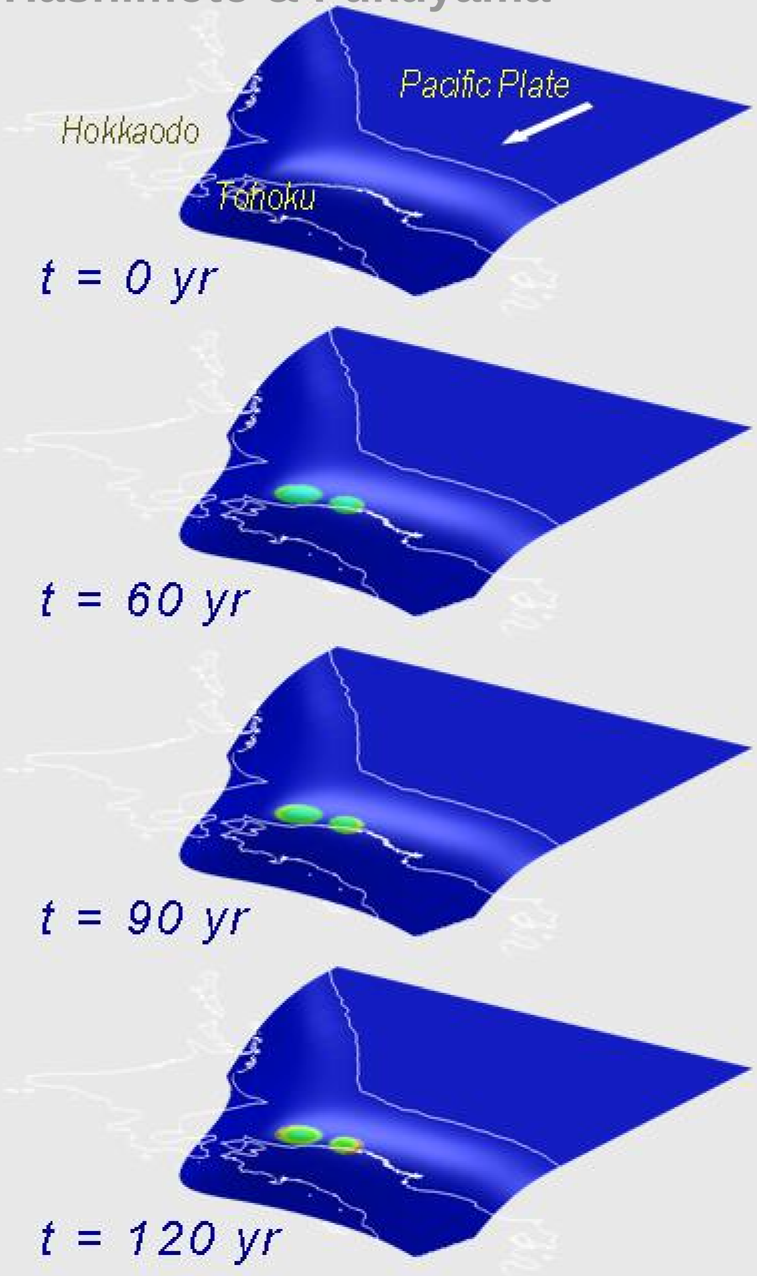
- Coupled Simulations: Current Status
  - In the previous projects in the Earth Simulator, each simulation model/code has been developed/updated, optimized for the Earth Simulator, parallelized for large-scale simulation.
  - We are now "coupling" them.

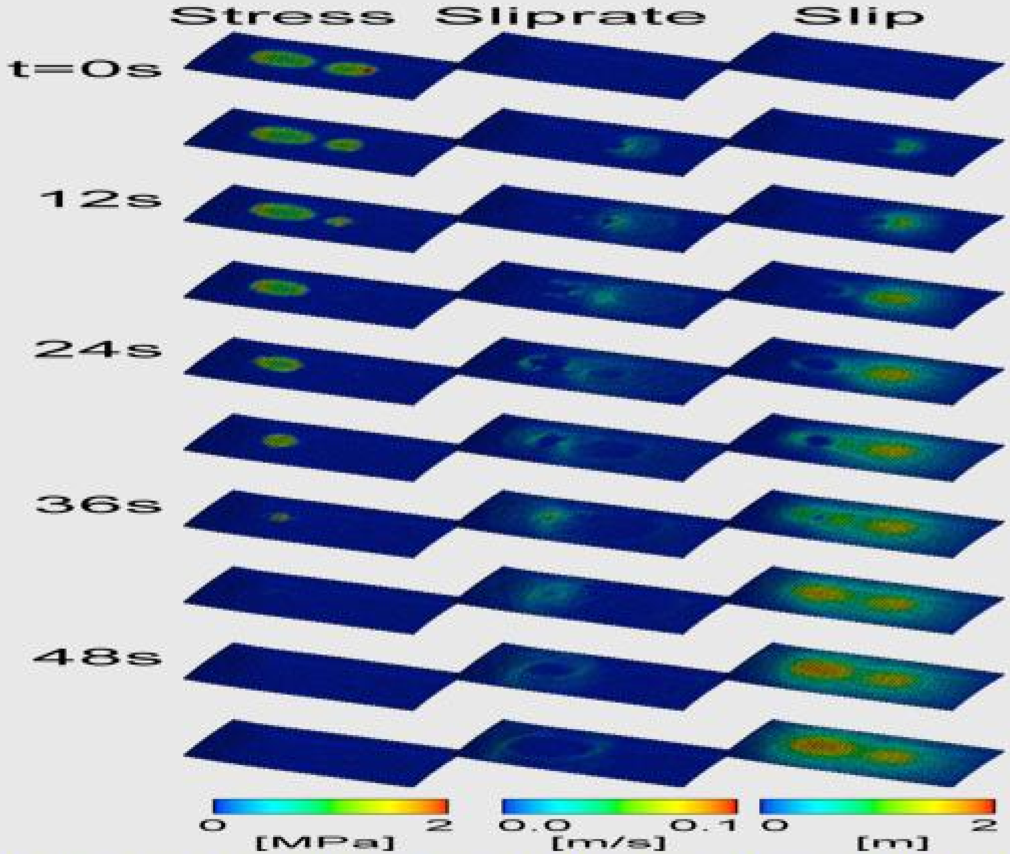
Outlines of the Integrated Predictive Simulation System



Hashimoto & Fukuyama

Joint Simulation of Earthquake Generation at the 1968 Tokati-oki Earthquake Seismogenic Region

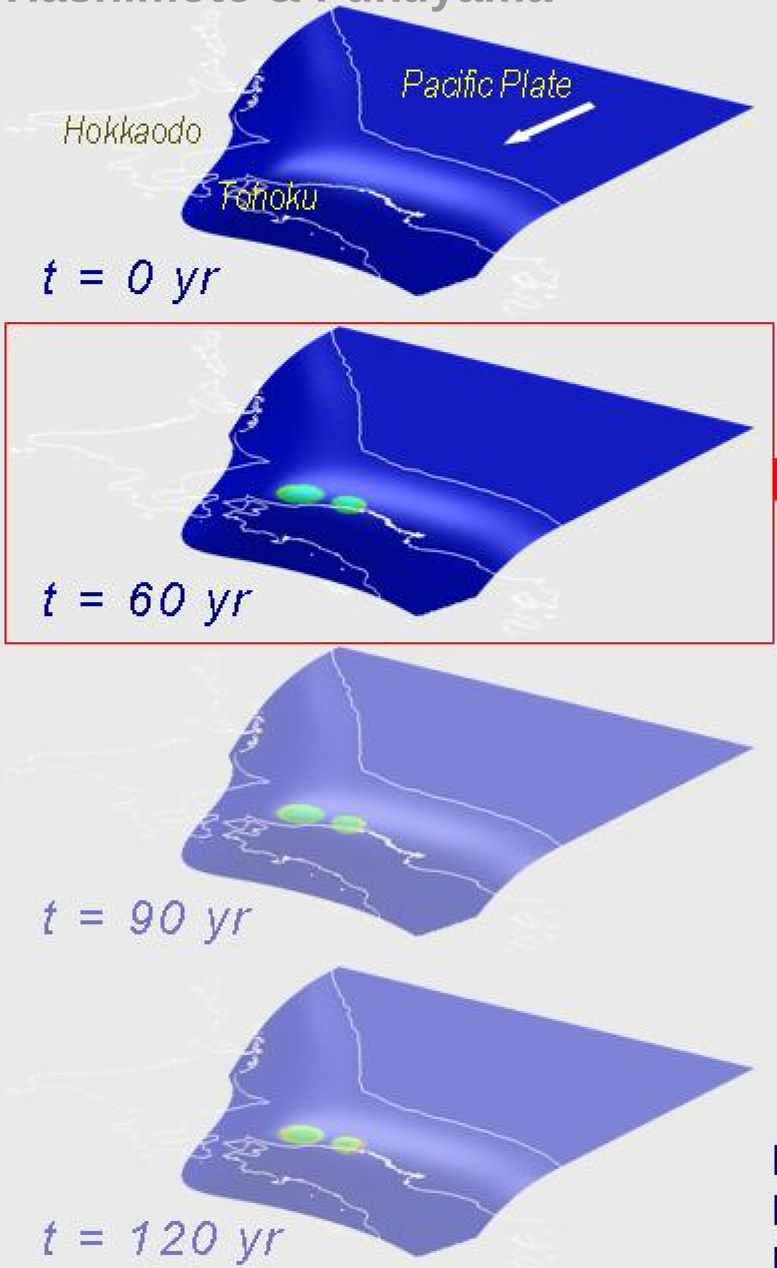


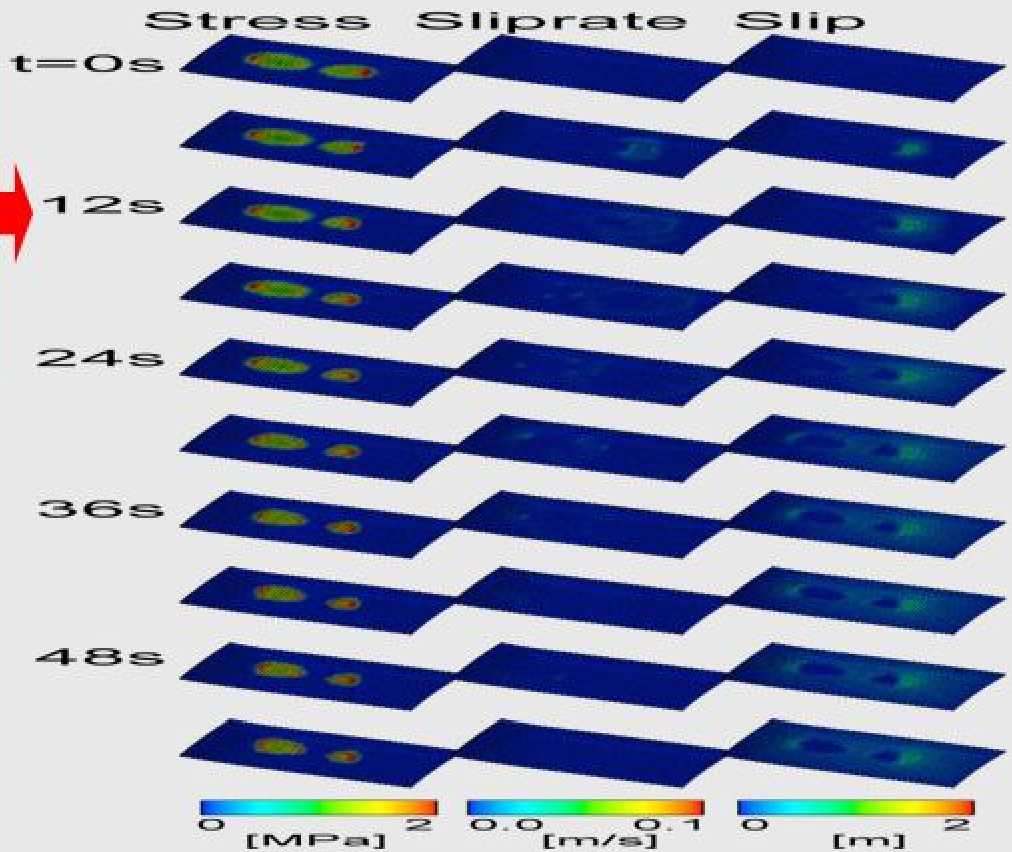


In the case where the stress level is close to a critical level, started dynamic rupture was accelerated rapidly and expanded over the whole seismogenic region.

Hashimoto & Fukuyama

Joint Simulation of Earthquake Generation at the 1968 Tokati-oki Earthquake Seismogenic Region

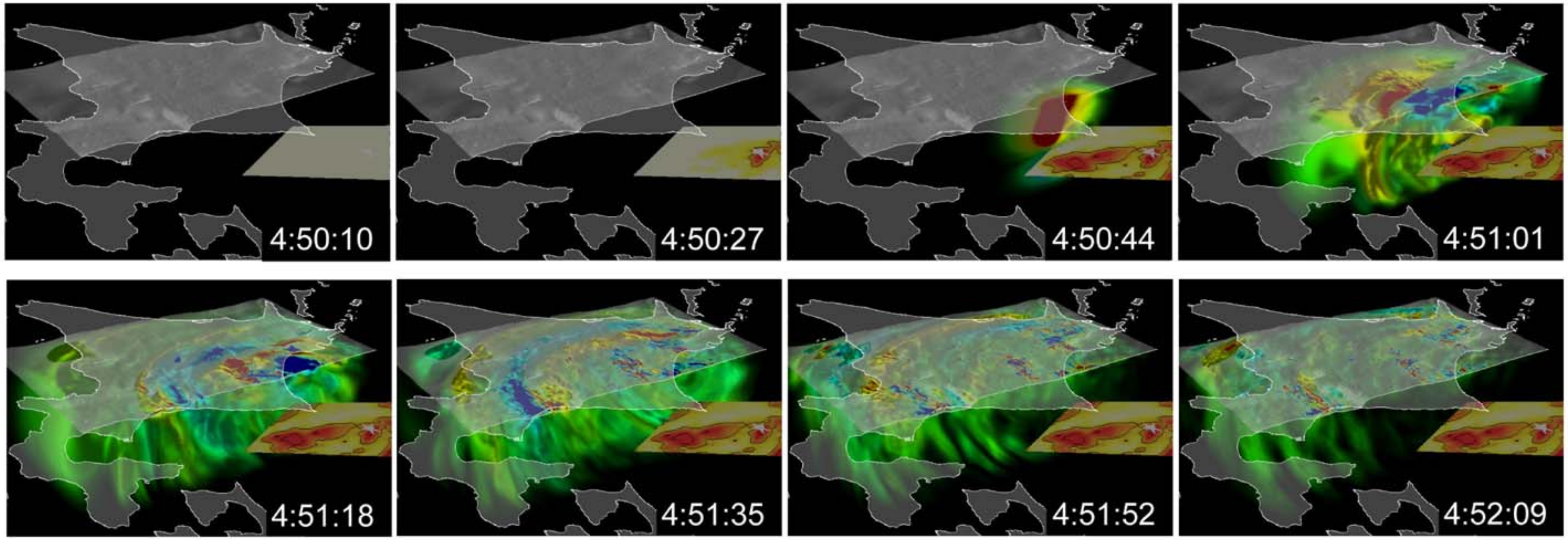


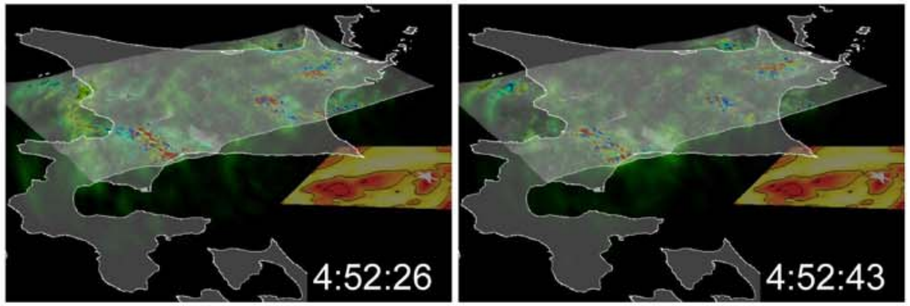


In the case where the stress level is not close to but not much lower than a critical level, started dynamic rupture was accelerated but expanded to only a part of the seismogenic region.

Joint Simulation of Dynamic Rupture and Seismic Wave Propagation

Computation of the seismic wave propagation associated with the dynamic rupture of the 2003 Tokachi-oki earthquake.



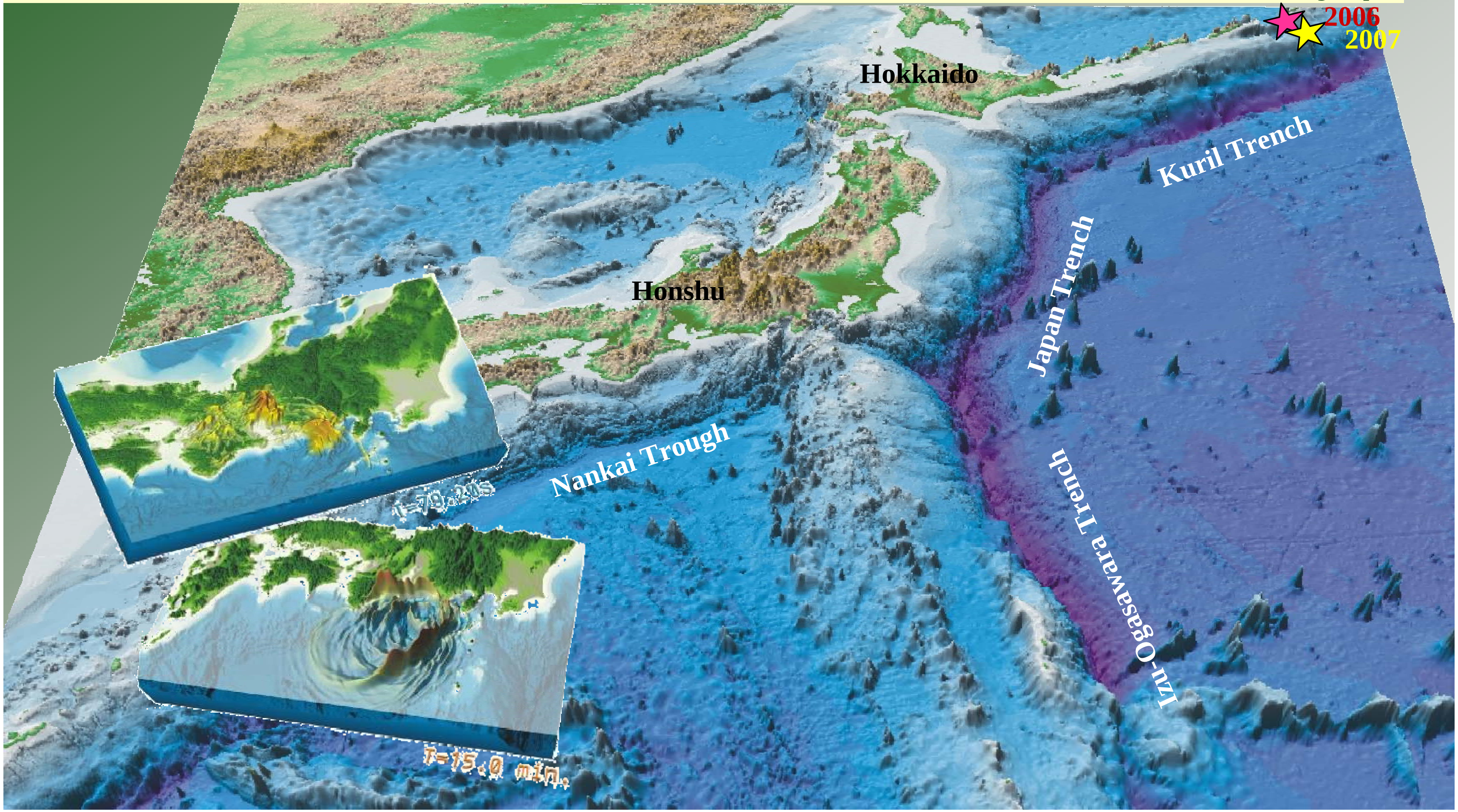


Snapshots of dynamic rupture and seismic wave propagation at 17 seconds intervals.

Fukuyama & Aoi

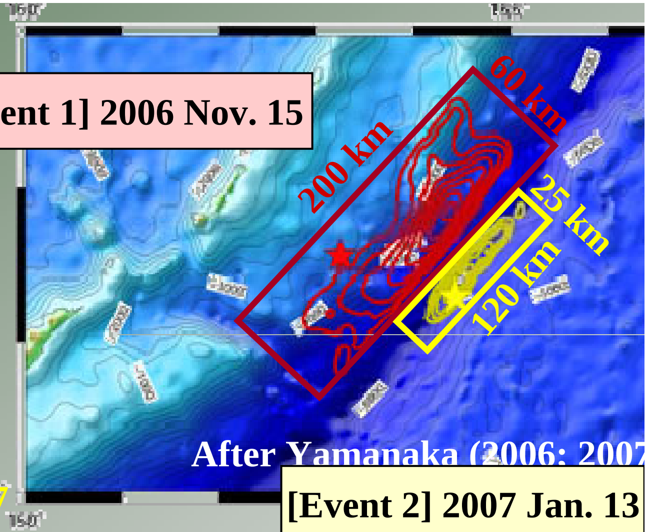
An Integrated Simulation of Seismic Wave and Tsunami Propagation

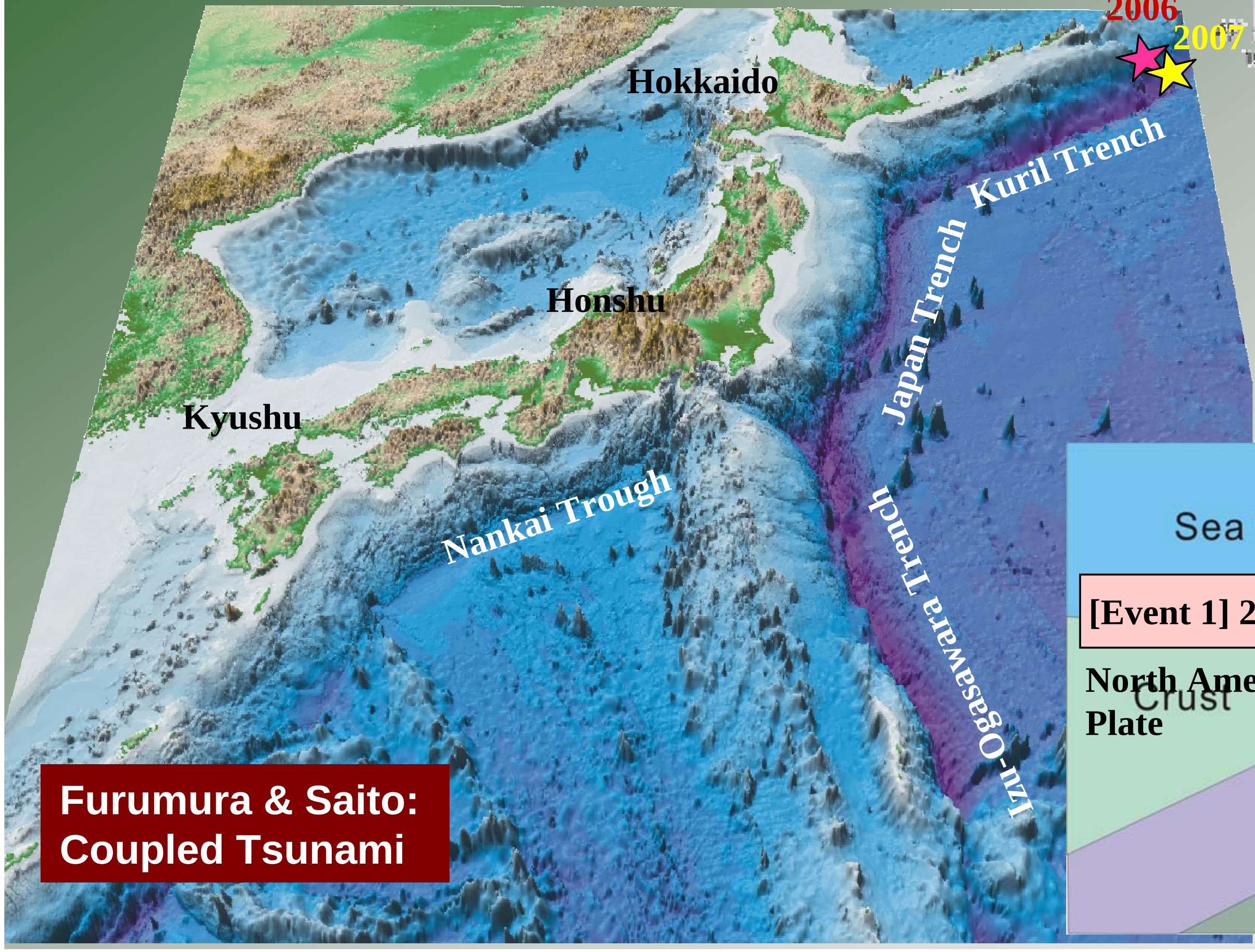
古村孝志・齊藤竜彦 (東大地震研)
Takashi Furumura & Tatsuhiko Saito （ERI. Univ. Tokyo)

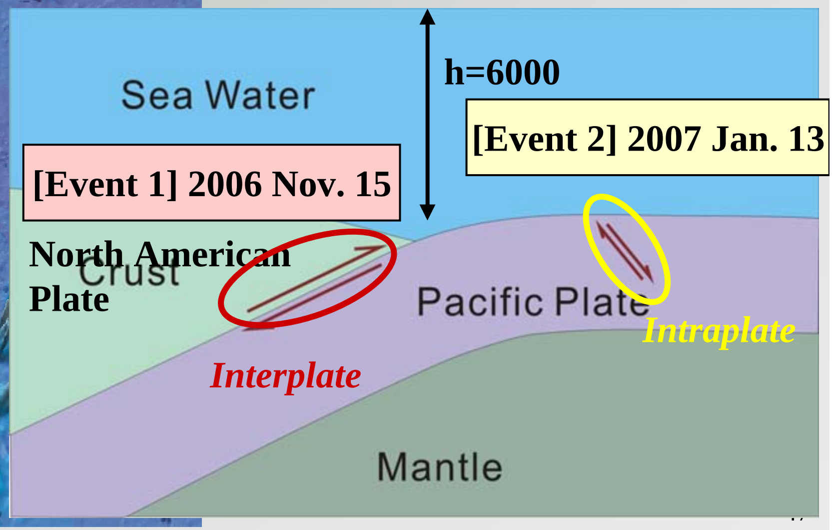


Characteristics of two Tsunami Events in Kuril Islands

Large M8 earthquakes occurred in Kuril Trench in 2006 and 2007; the former is an interplate event and the other is an intraplate event



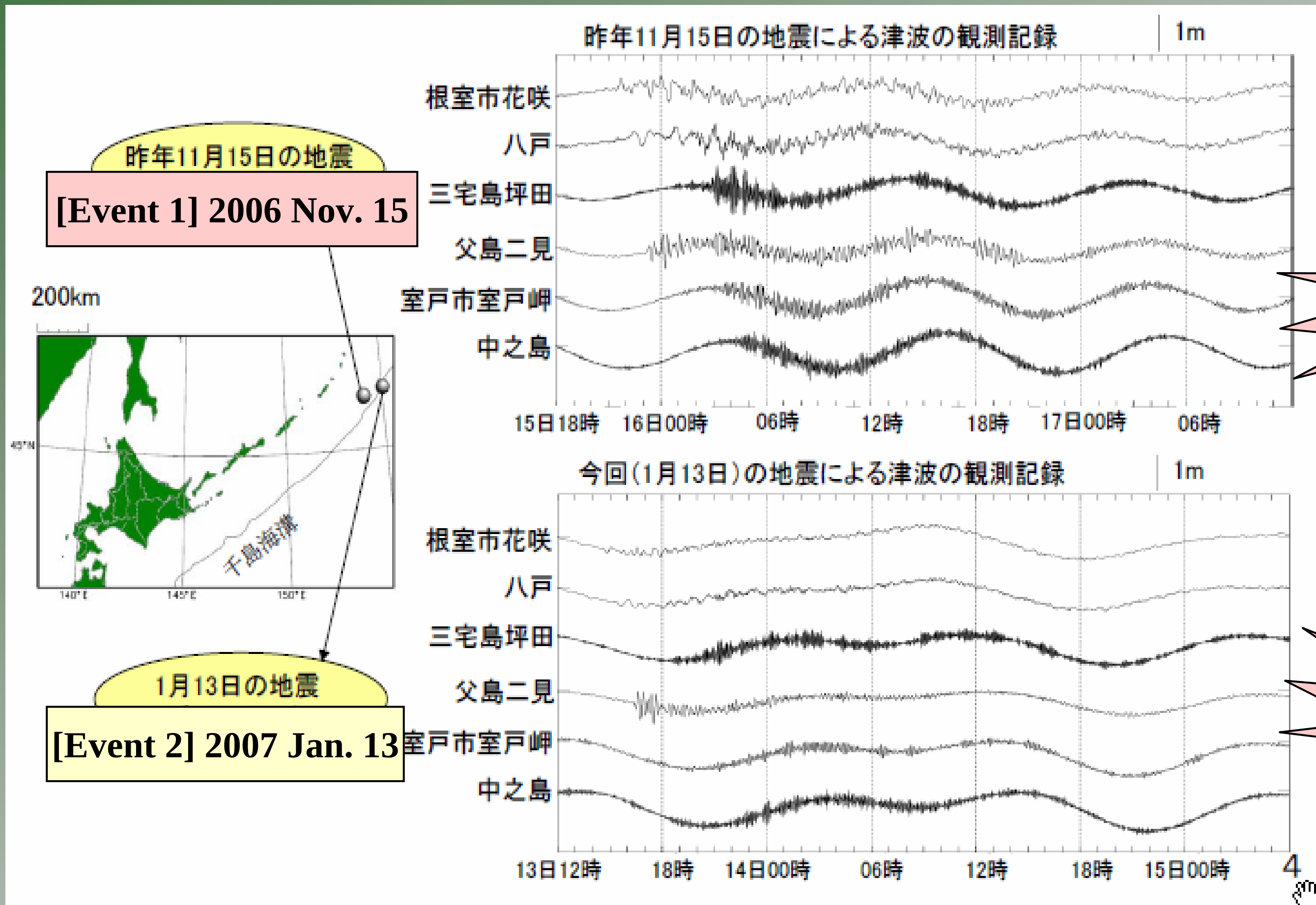




Furumura & Saito: Coupled Tsunami

Observed Tsunami

Tide gage record shows larger tsunami from the 1st (2006) event and very weak tsunami from the 2nd (2007) event.



Observation:

Hachinohe: 53cm

Under Estimation

Hachinohe: 17cm

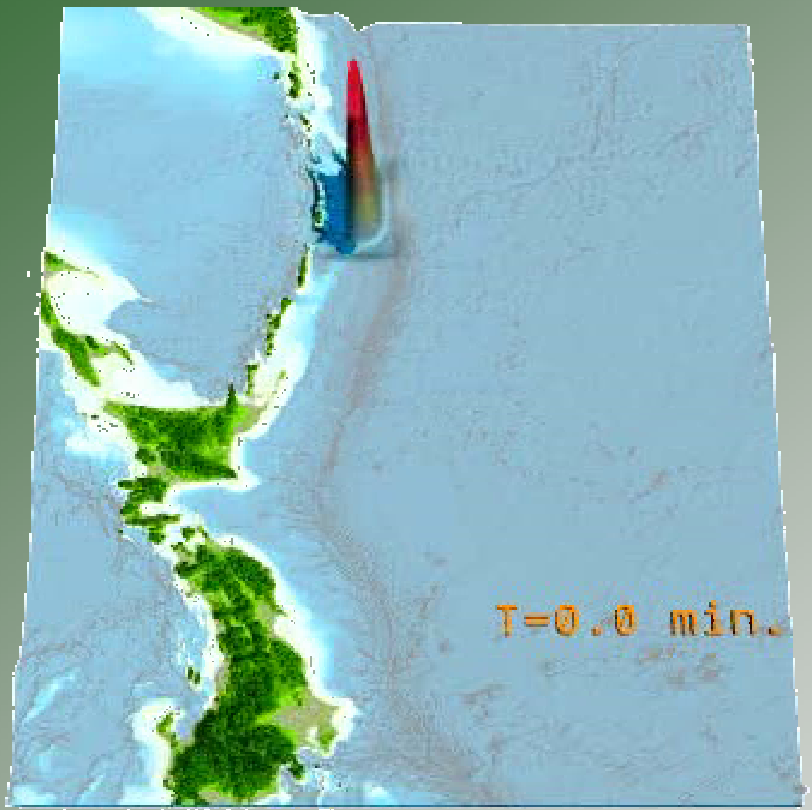
Over Estimation
Mistake Alert !

After JMA (2007)

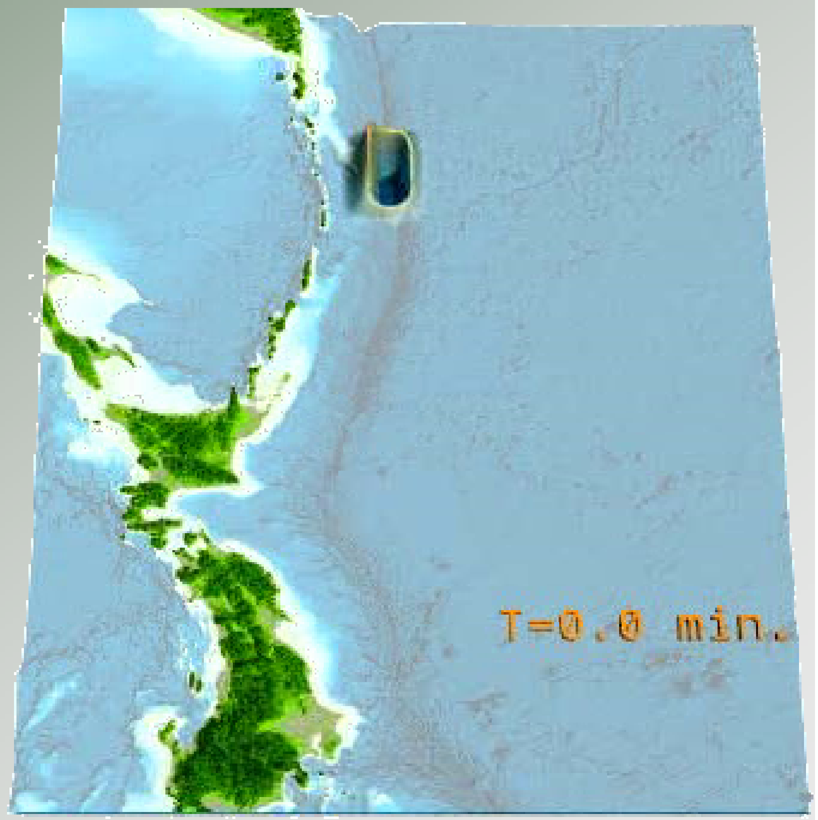
Tsunami Simulation
Traditional Shallow Water Eqn’s

Furumura & Saito: Coupled Tsunami

[Event 1] 2006 Nov. 15



[Event 2] 2007 Jan. 13

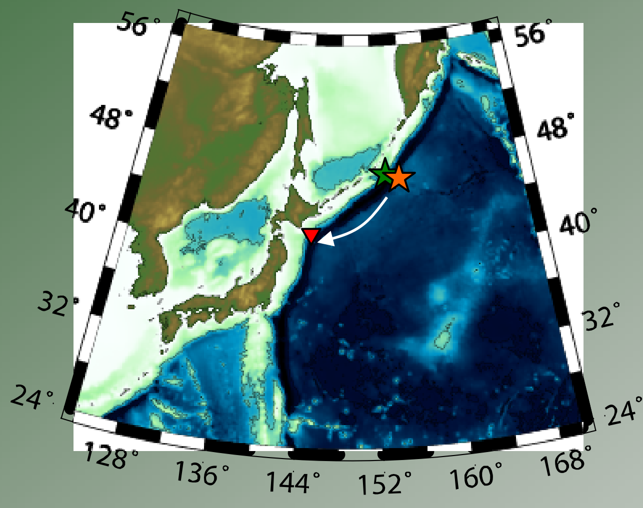


A parallel tsunami simulation code (Saito and Furumura, 2007) is used which took 30 min using 16CPU of AMD Opteron.

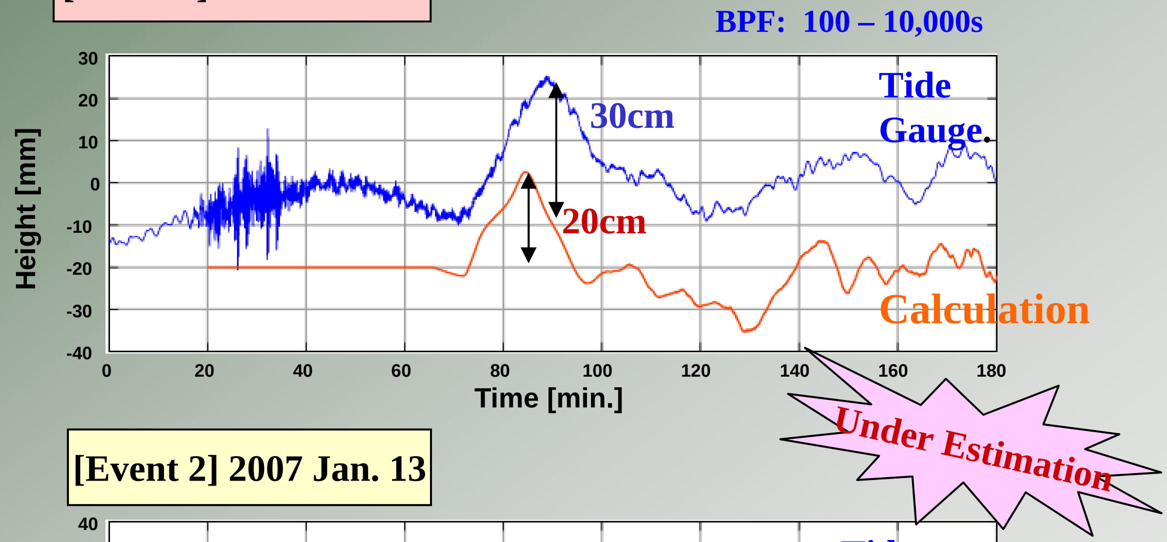
Simulation Results

Furumura & Saito: Coupled Tsunami

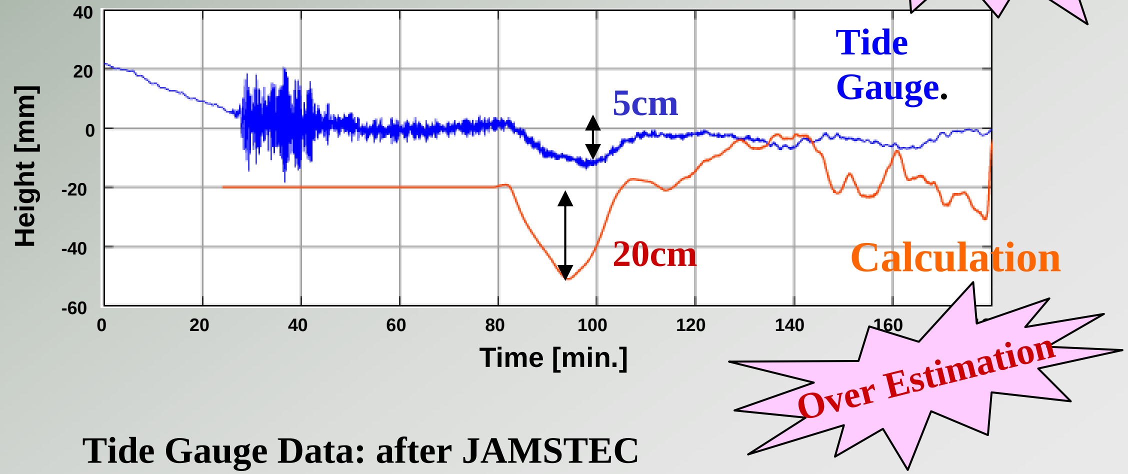
Simulation results are compared with the tide gauge data at offshore Tokachi. It is indicating under and overestimation of tsunami for 1st and 2nd events, respectively, – similar to JMA alert.



[Event 1] 2006 Nov. 15

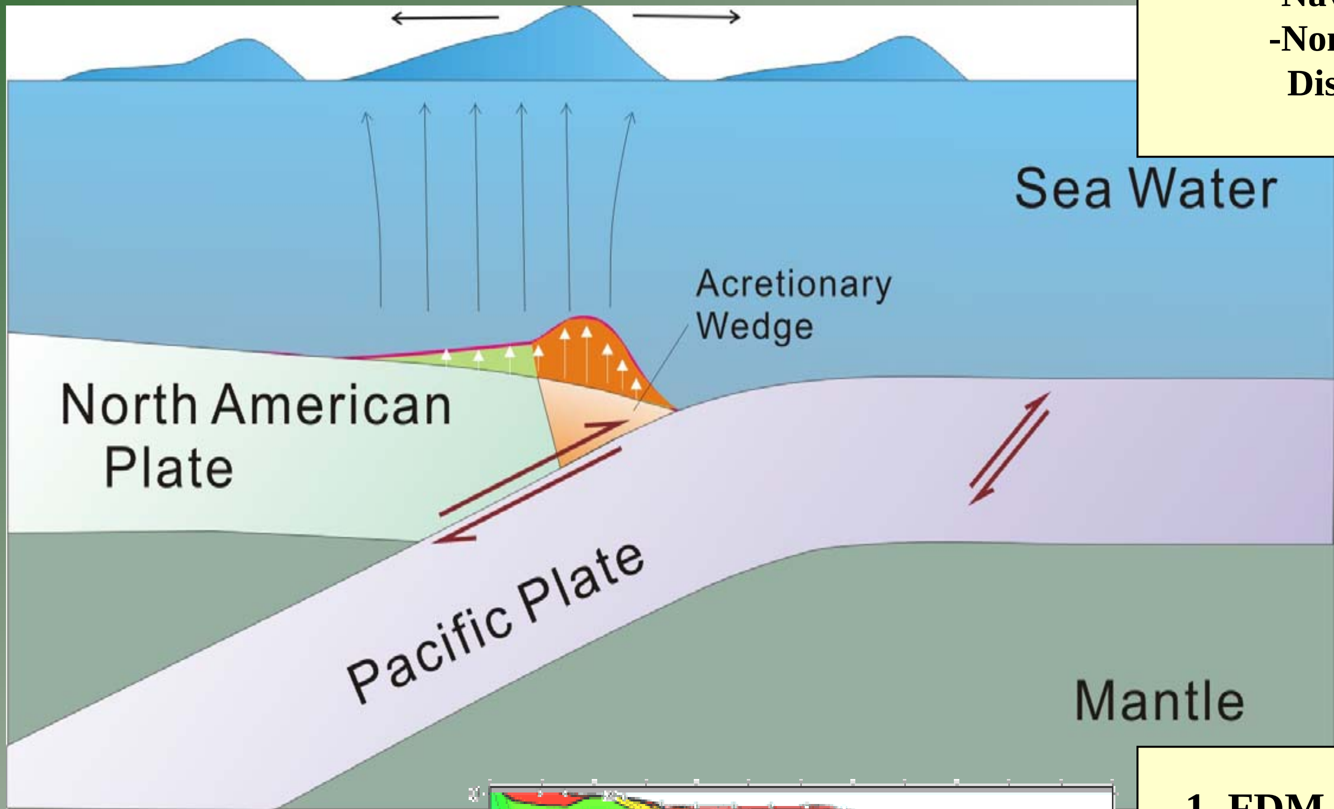


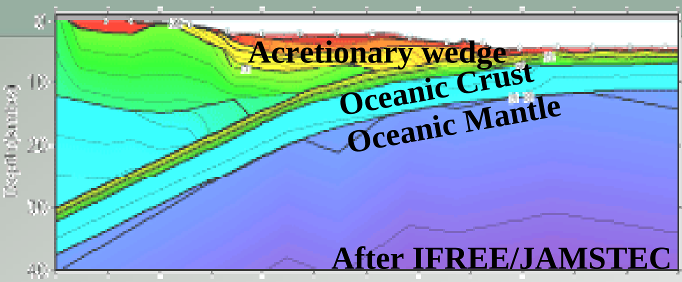
[Event 2] 2007 Jan. 13



Tide Gauge Data: after JAMSTEC

Accurate Tsunami Simulation—Challenge





Furumura & Saito: Coupled Tsunami

2. FDM Simulation of tsunami generation/propagation

- **-Navier-Stokkes Equations in 3D**
- **-Nonlinearity, Viscosity Friction, Dispersion, etc**

V (x,y,t) or P (x,y,t)

Coupling (one way)

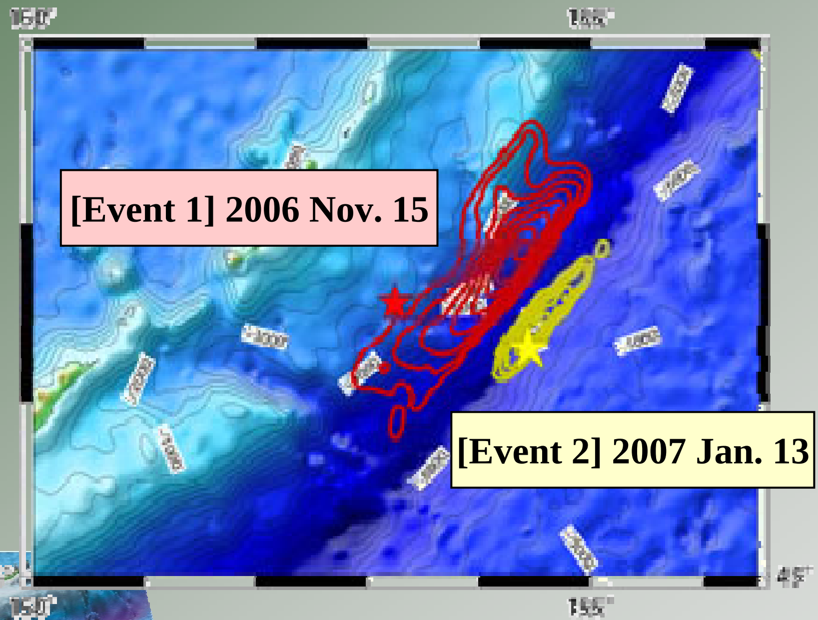
V (x,y,t)

1. FDM Simulation of Seismic Waves

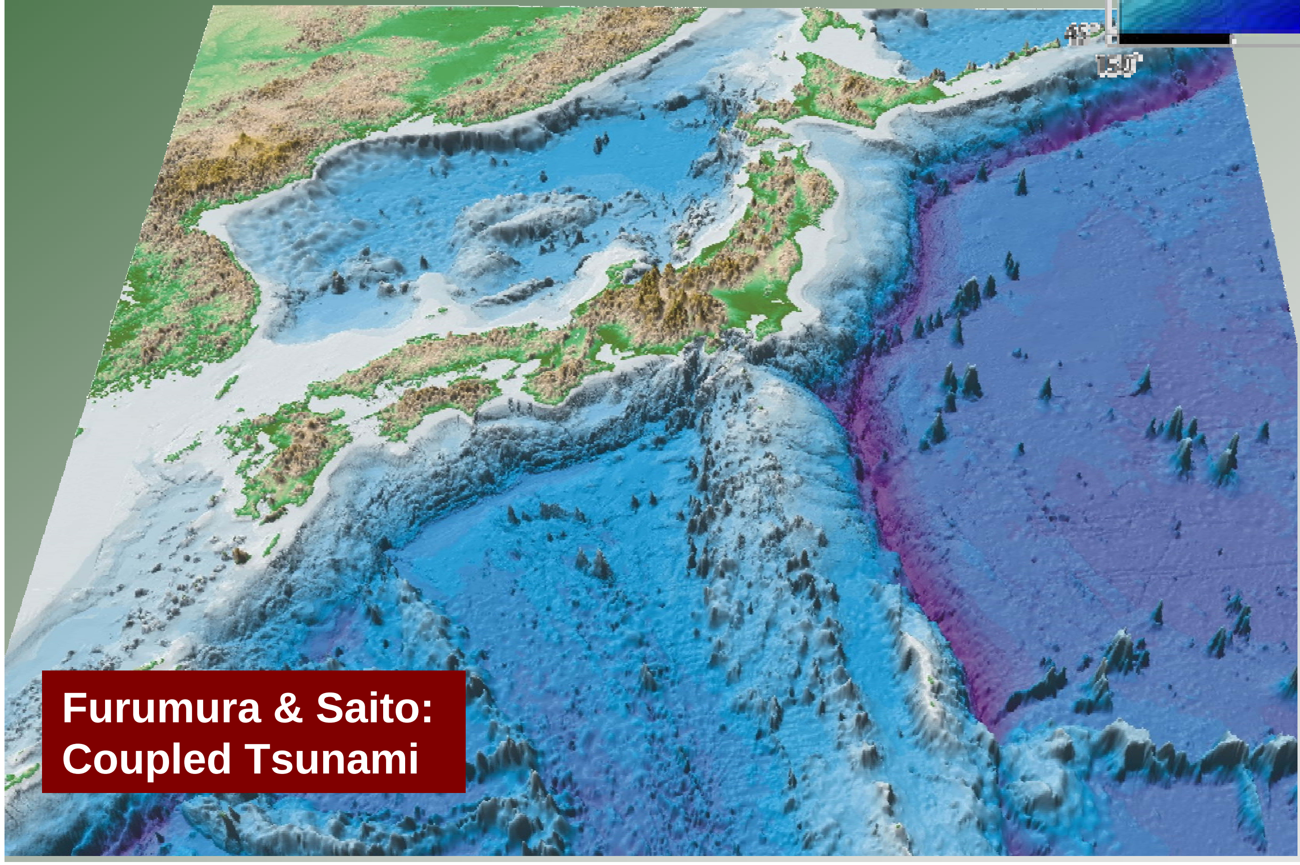
- **- Equation of Motions in 3D**
- **- 3D Heterogeneous structure**
- **- Source Slip model**

Effect of Deep Sea

The Long-wave, shallow water approximation used in the conventional tsunami simulation does not simulate tsunami propagation in deep (6000-8000m) sea ?



Sea Depth: 6000-8000m



Furumura & Saito: Coupled Tsunami

Full 3D FDM Simulation of Tsunami

Direct tsunami simulation without approximations

- Mass continuity equation (incompressible flow)

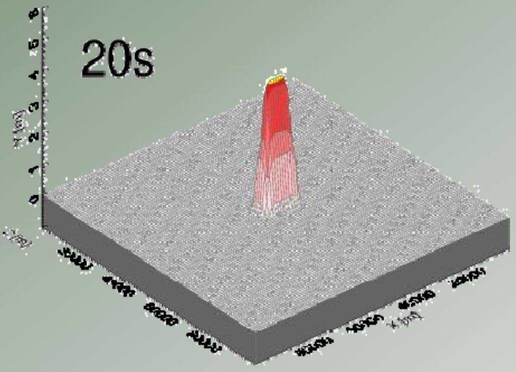
$$\nabla \cdot \mathbf{u} = 0$$

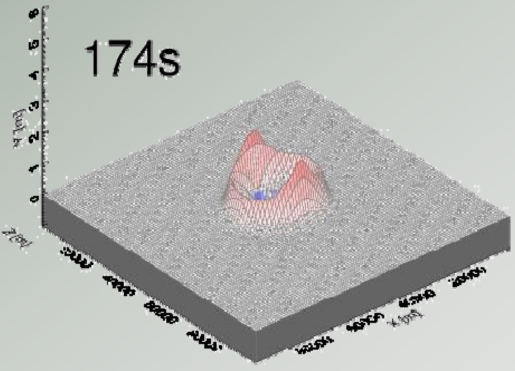
$\mathbf{u} = (u, v, w)$: velocity vector

- Navier-Stokes Equation

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \nu \Delta \mathbf{u} - \mathbf{g}$$

p: pressure, ν: kinematic viscosity coefficient, $\boldsymbol{g}$: gravity vector





- Boundary Conditions

Free surface at the top

$$\frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} = w$$

$h\,(x,y)$: height of the top surface

Pressure at the top

$$p\,(x, y, z = h) = 0$$

Rigid boundary at the bottom

$$u_n = 0$$

- 3D FDM simulation of NS

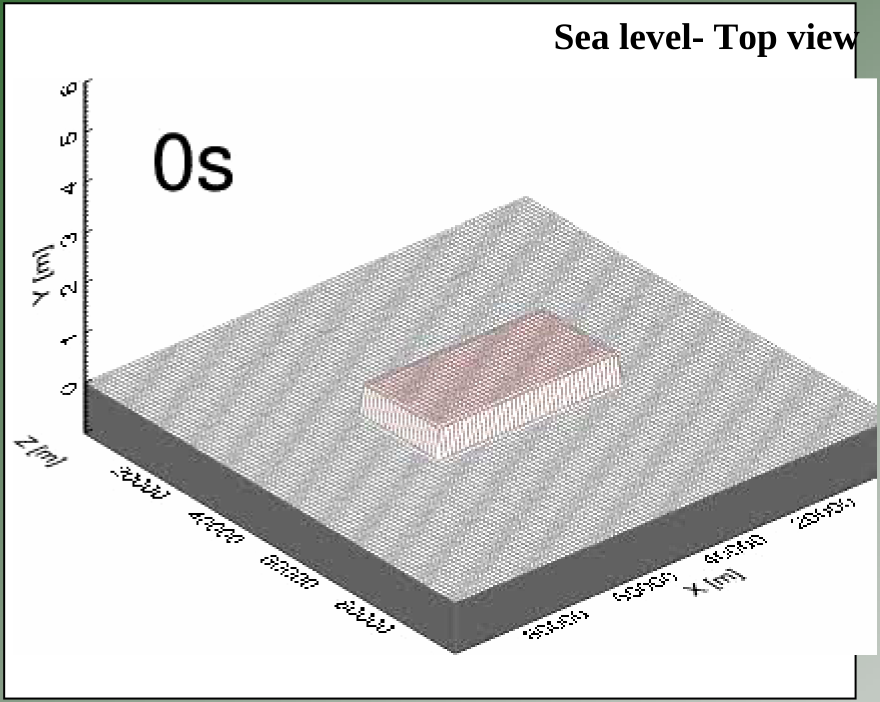
SOLA-SURF in 3D
(e.g. Hirt et al. 1975, LLNL)

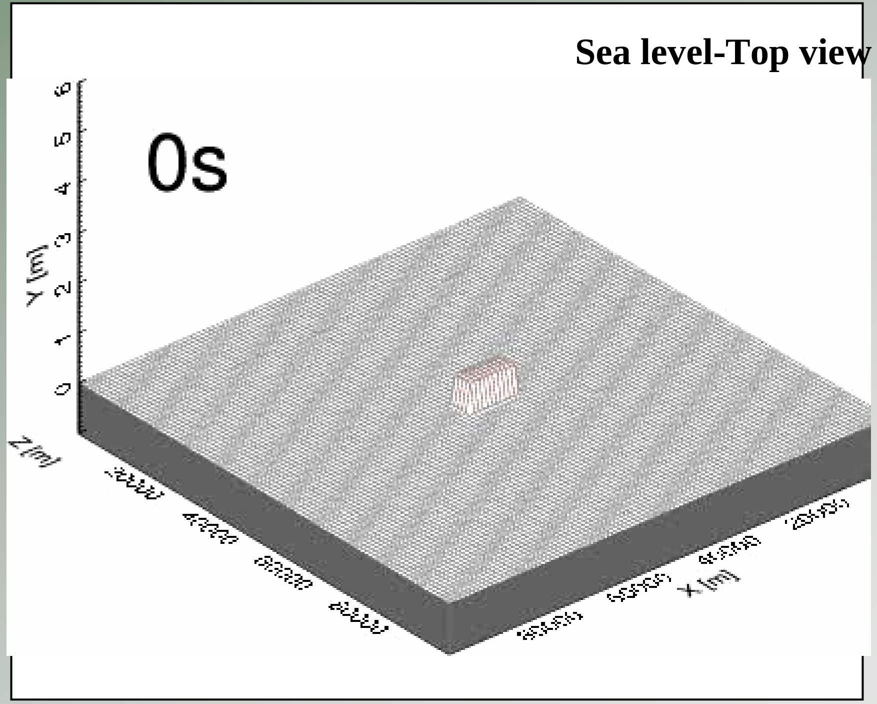
Numerical simulation of tsunami generation: (1) Shallow (1000m) water

Furumura & Saito: Coupled Tsunami

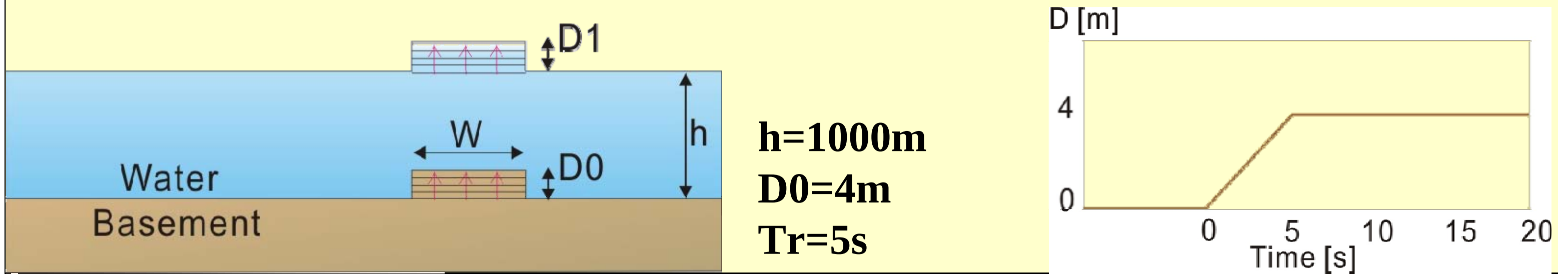
(a) Large Fault (W/L=20km/40km)

(b) Small Fault (W/L=10km/5km)





Vertical Deformation of Seafloor



h=1000m
D0=4m
Tr=5s

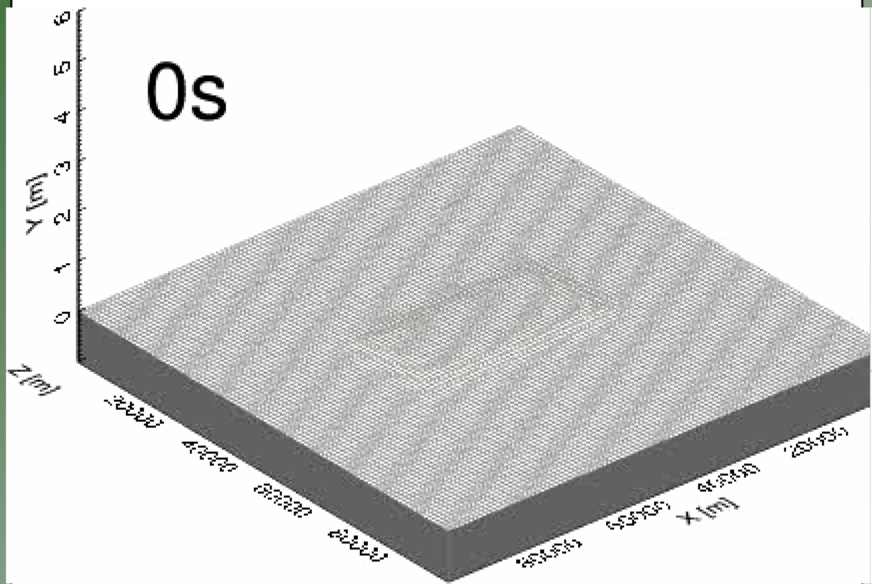
Numerical simulation of tsunami generation: (2) Deep (6000m) water

Furumura & Saito: Coupled Tsunami

Thick water column cannot push up sea level very efficiently, and so the Initial tsunami height is much lower than the vertical deformation of seabed

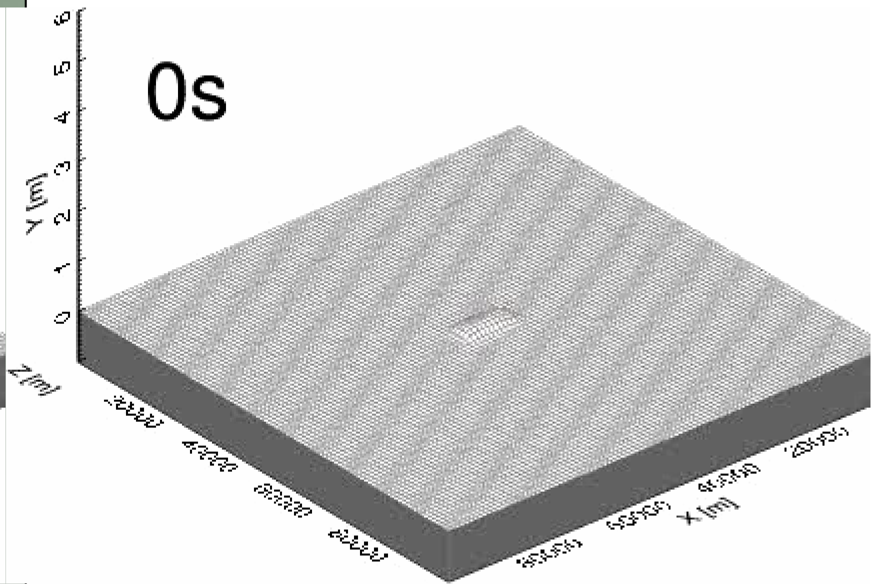
(a) Large Fault (W/L=20km/40km)

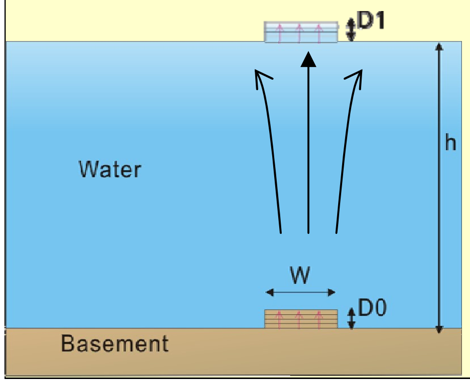
Sea level



(b) Small Fault (W/L=10km/5km)

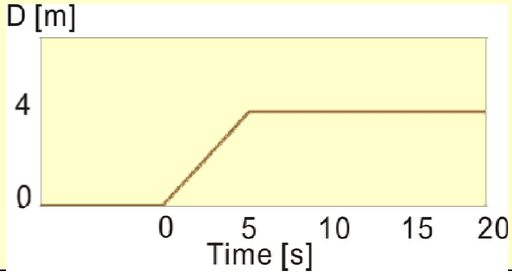
Sea level





Deep Sea (h=6000m)
D0=5m
W=20km, 60km

Vertical Deformation of Seafloor



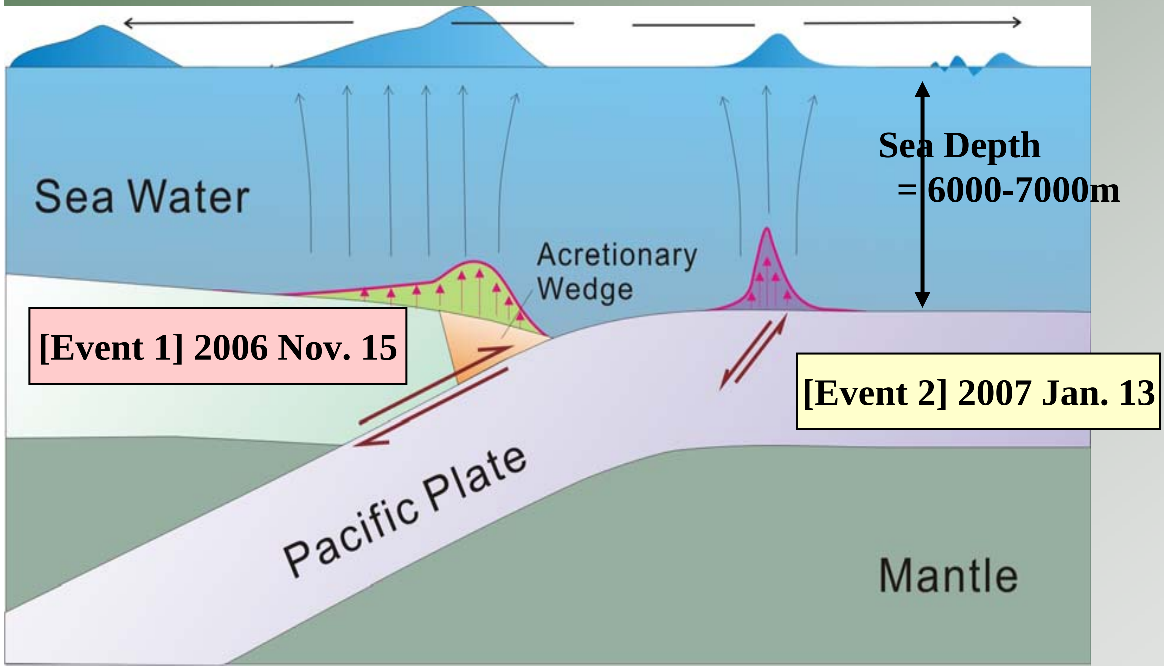
Tsunami Simulation - Summary

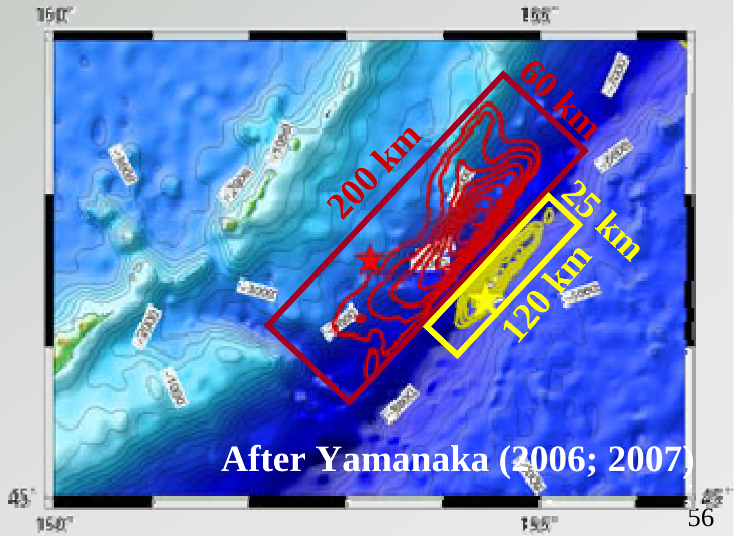
Furumura & Saito: Coupled Tsunami

	[Event 1] 2006 Nov. 15 Mw8.2	[Event 2] 2007 Jan. 13 Mw8.2
Fault Size (L*W)	Large: 200km*60km	Small: 25km*120km
Large Deformation in Acretionary Wedge	may be	no
Sea Depth	Deep: >6000m	Deep: >6000m
Push up Sea surface	Efficient	Not efficient
Attenuation by Dispersion	Weak	Strong

Larger Tsunami

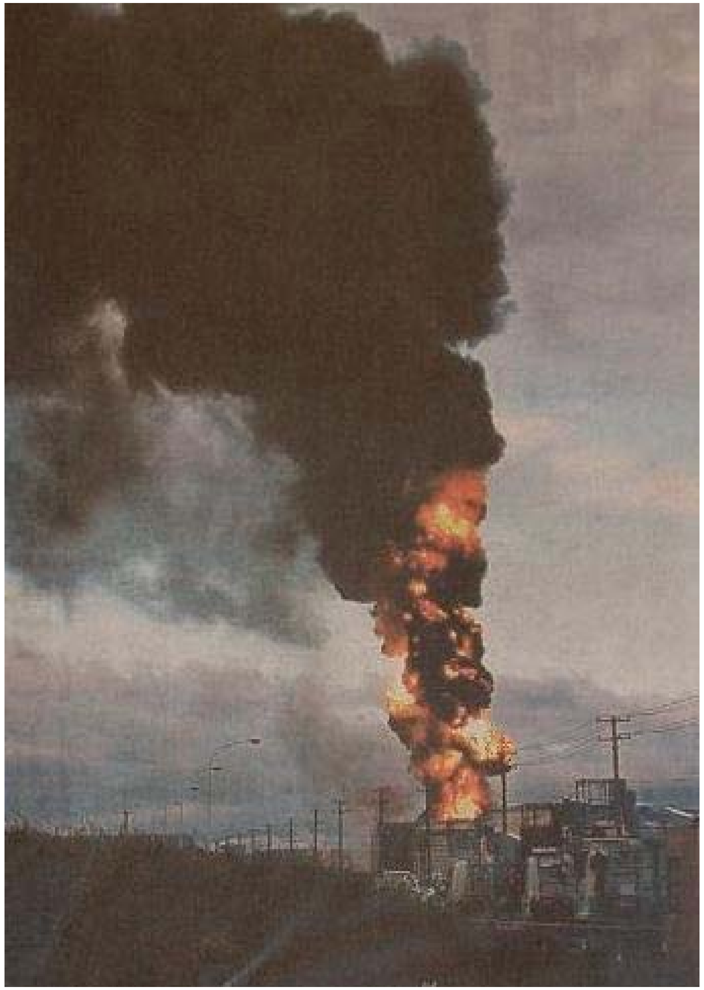
Weak Tsunami

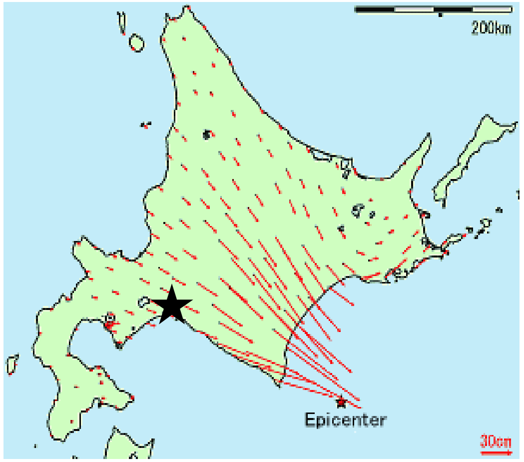




2003 Tokachi Earthquake (M8.0)

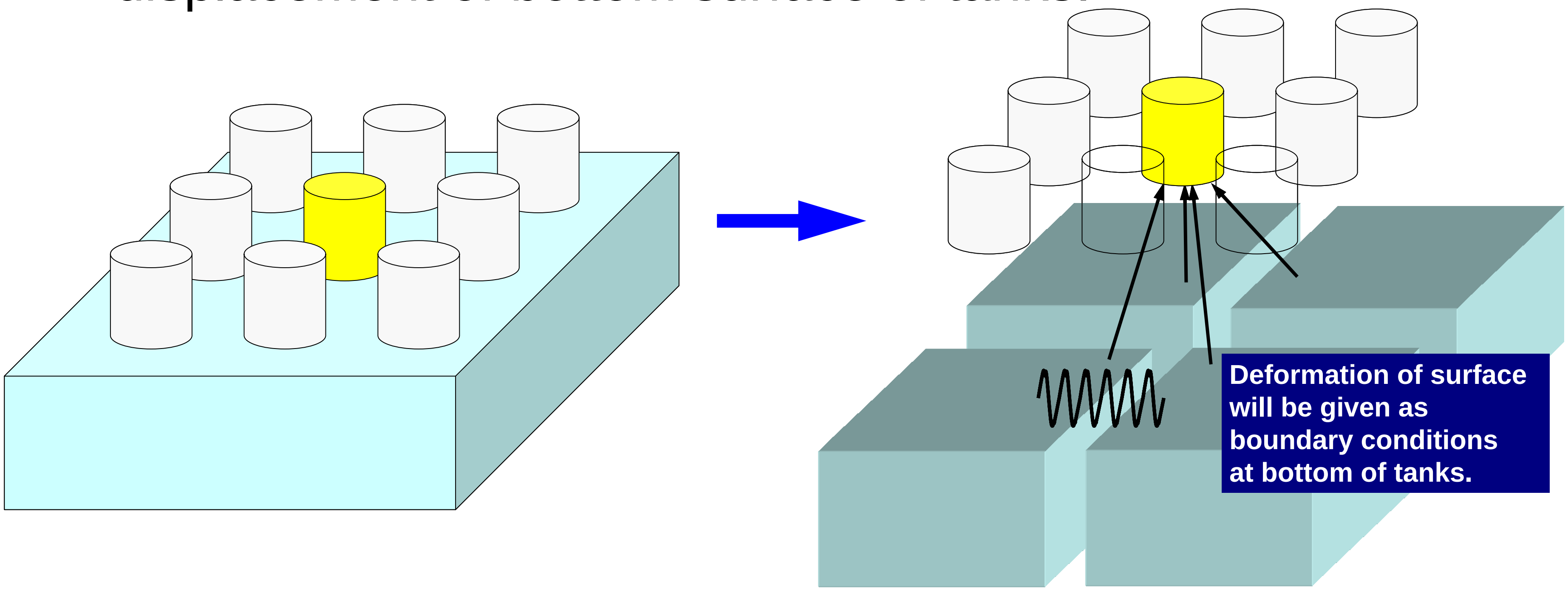
Fire accident of oil tanks due to long period ground motion (surface waves) developed in the basin of Tomakomai





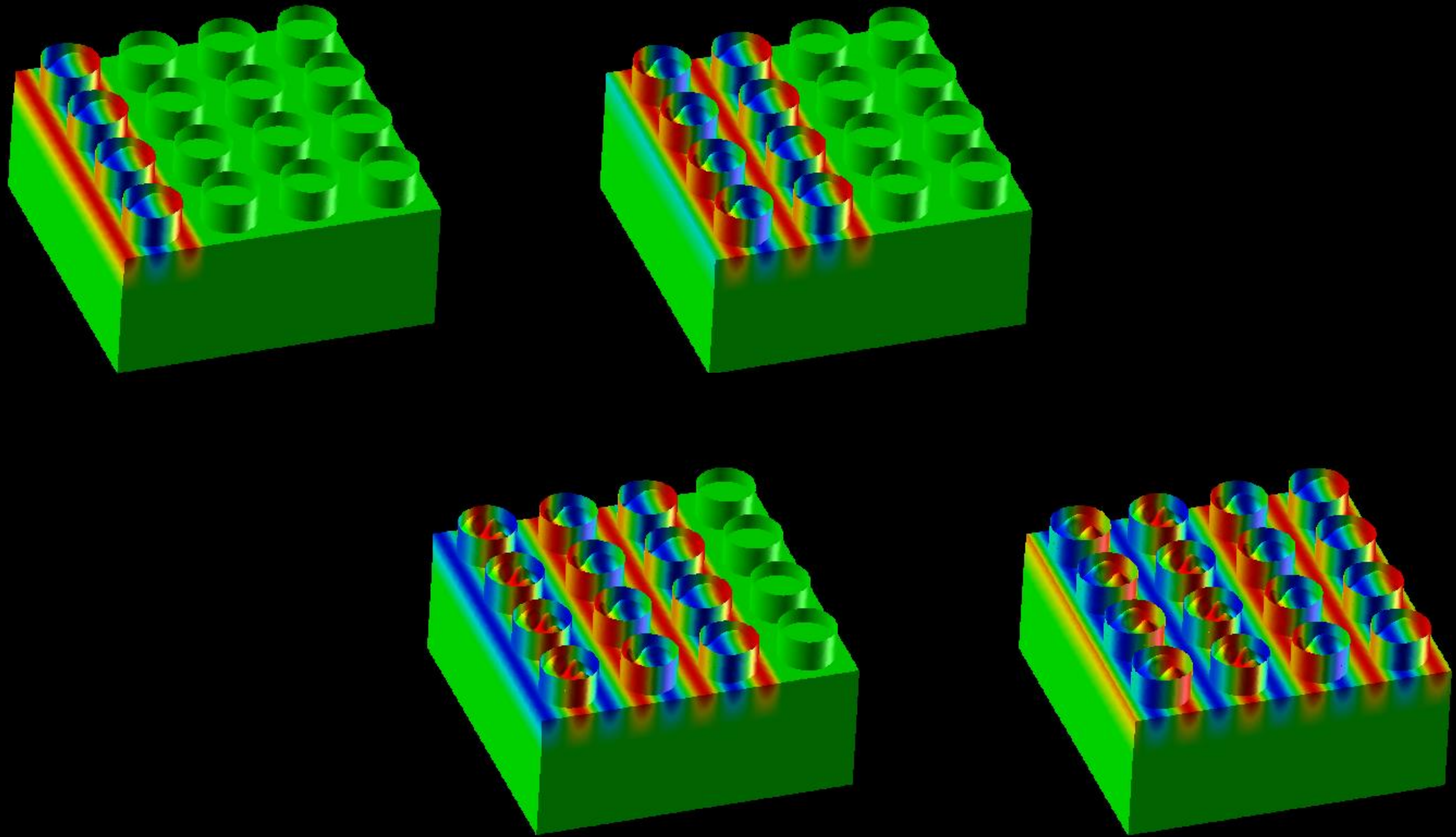
(Current) Target Application

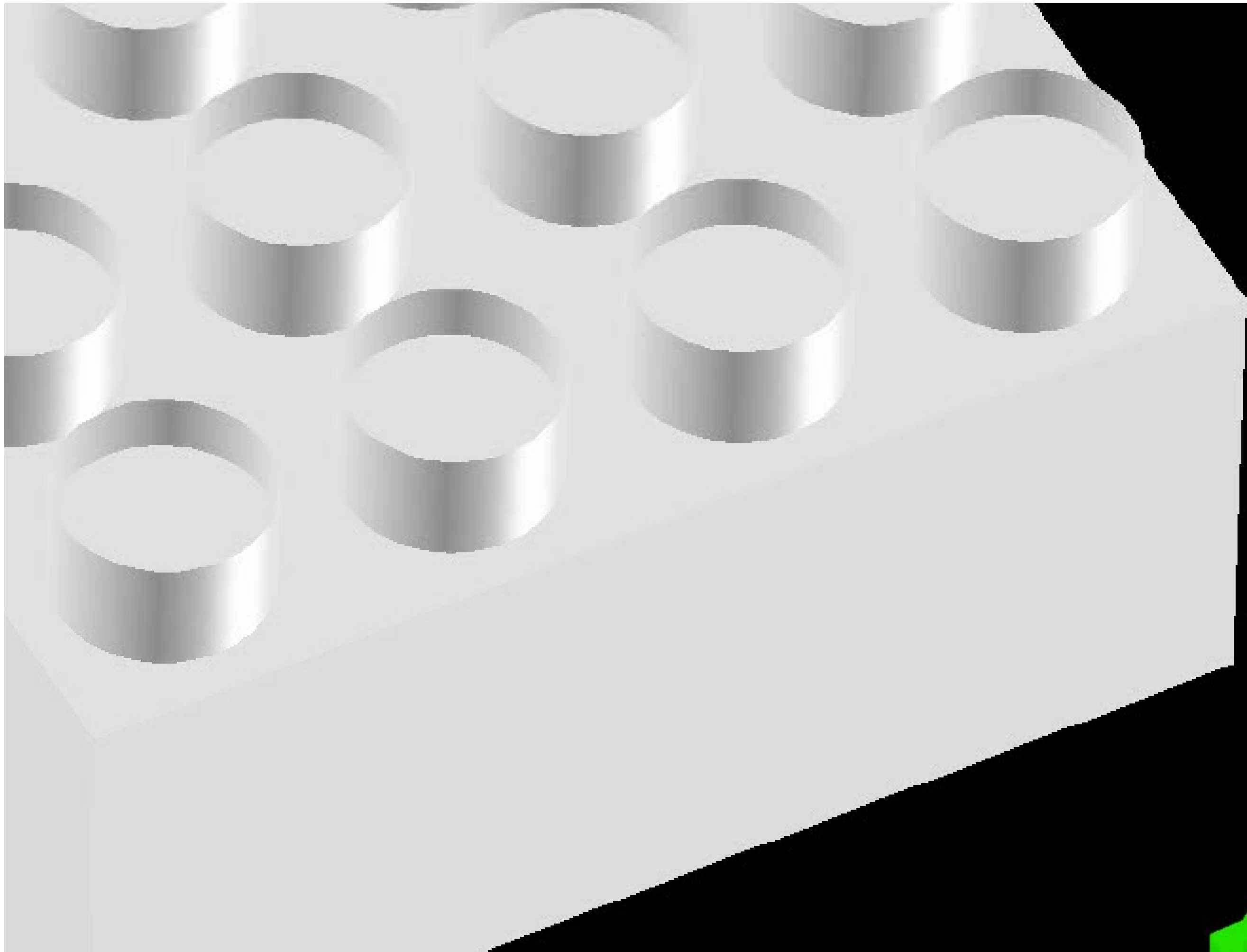
- Coupling between "Ground Motion" and "Tanks for Oil-Storage"
  - "One-way" coupling from "Ground Motion" to "Tanks".
  - Each of appl. knows "number" of processes of the other one.
  - Displacement of ground surface is given as forced displacement of bottom surface of tanks.

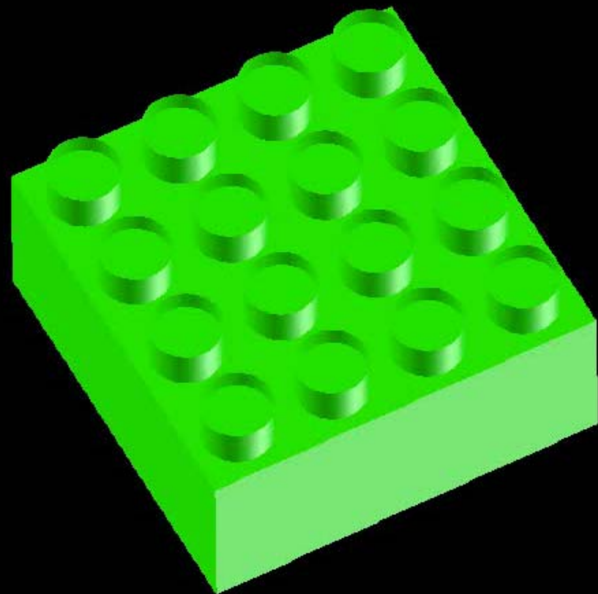


Parallel Simulations using 32 cores

16 for tanks, 16 for ground motion







Parallel Visualization by AVS/Express PCE

- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- Overview of the Current Project by JST
  - Integrated Predictive Simulation System for Earthquake and Tsunami Disaster
- **Some Technical Issues**
  - **Parallel Preconditioning Methods**
  - Vector vs. Scalar Processors
  - Parallel Programming Models in Multi-Core Era
- Future Directions

Geophysics and Computer/Computational Sciences

- *Need to engage computational scientists in directing their research to solve our computational issues.*

  - *Brad Aagaard (June 26, 2007)*

- MY motivations and "research cycle"
  - to find solutions for specific (and difficult) problems (e.g. geophysics).
  - to construct general algorithms, to develop general frameworks.
  - to apply the results to more general problems.
- One successful example

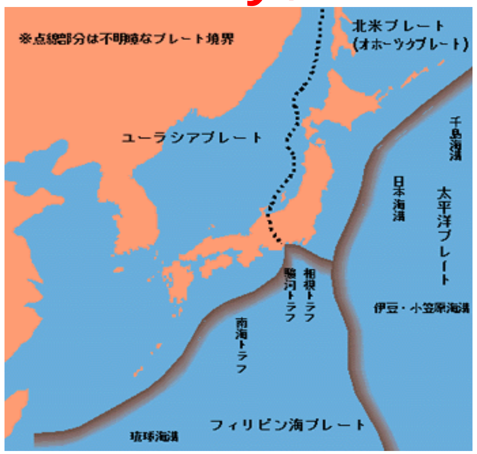
  - special preconditioning method for fault contact problems
  - extension to general applications

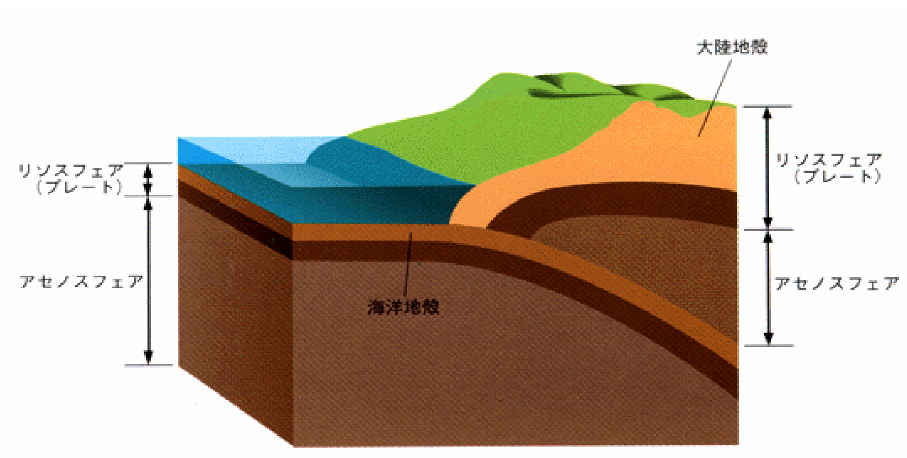
Overview: Contact Problems

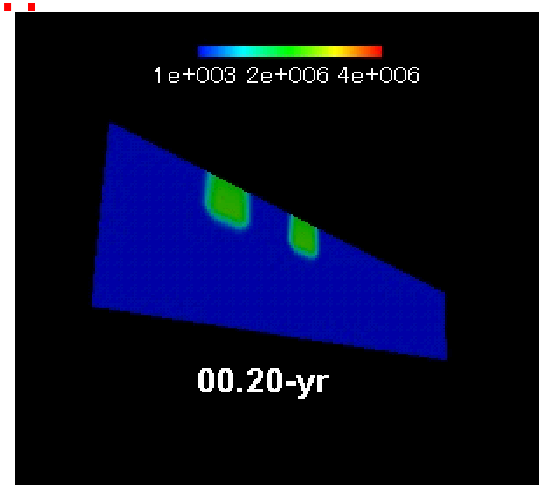
- Background
  - Simulations for Earthquake Generation Cycle
  - Selective Blocking
- More General Problems
  - Extension of Overlapped Zones
- Preconditioning/Partitioning Methods
  - Target Application
  - Selective Fill-in
  - Selective Overlapping
- Results
- Summary
  - Future Works

Contact Problems in Simulations of Earthquake Generation Cycle

- Quasi-static stress accumulation process at plate boundaries
- Non-linear contact problems with Newton-Raphson iter's
- Ill-conditioned linear equations due to penalty constraint by ALM (Augmented Lagrangean).
- Parallel FEM with domain decomposition (GeoFEM)
- Finally, we adopted dislocation approach…

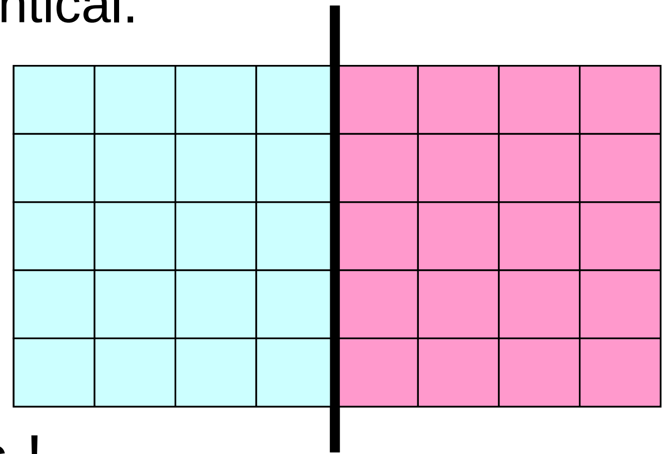






Contact Problems in Simulations of Earthquake Generation Cycle (cont.)

- Assumptions (GeoFEM)
  - Infinitesimal deformation, static contact relationship.
    - Location of nodes is in each "contact pair" is identical.
    - “Consistent” node number and position
  - No friction : Symmetric coefficient matrix



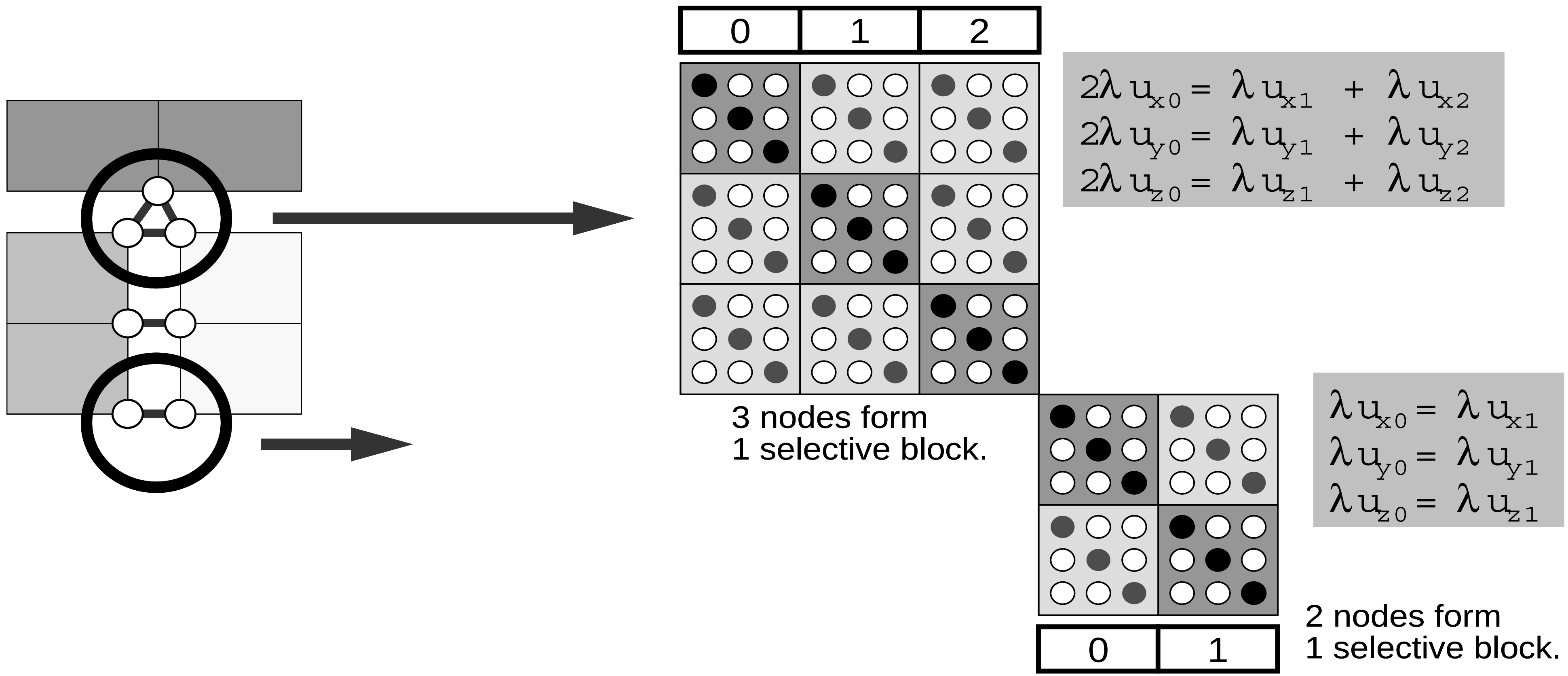
- Large-scale problems
  - Parallel preconditioned iterative solvers
  - We need robust & efficient preconditioners !
- Special preconditioning : ***Selective Blocking.***
  - provides robust and smooth convergence in 3D solid mechanics simulations for geophysics with contact.

Selective Blocking [Nakajima, 2001]

Special Method for Contact Problems

Strongly coupled nodes are put into the same diagonal block.

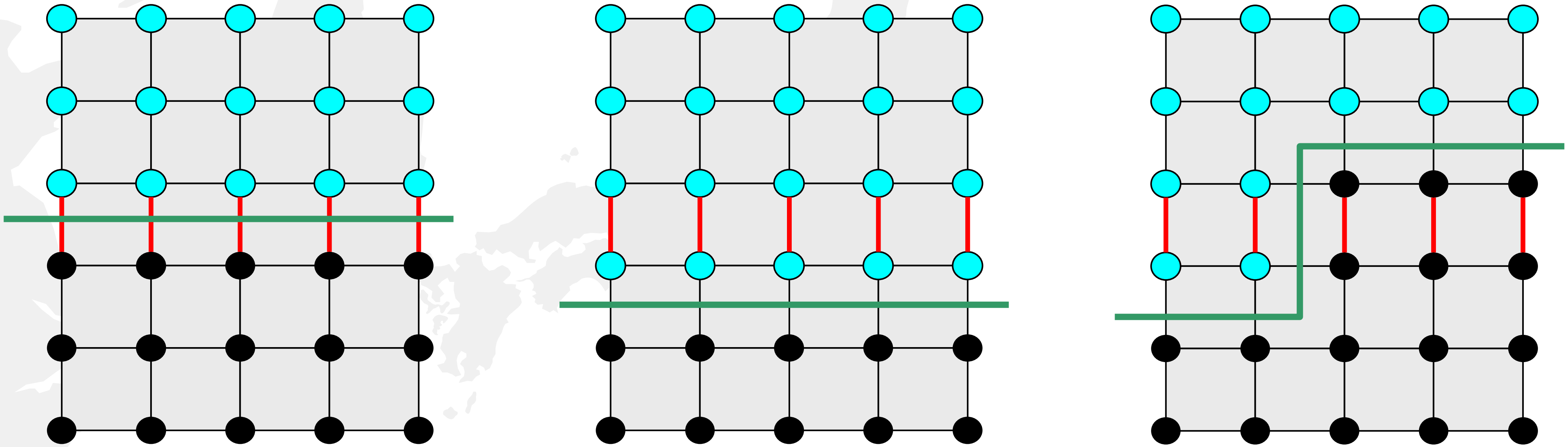
Full LU factorization for each block.



Special Partitioning Method

Convergence is slow if nodes in each contact group locate on different partition.

Repartitioning so that nodes in contact pairs would be in same partition as INTERIOR nodes will be effective.



BEFORE repartitioning

Nodes in contact pairs are on separated partition.

AFTER repartitioning

Nodes in contact pairs are on same partition, but no load-balancing.

AFTER load-balancing

Nodes in contact pairs are on same partition, and load-balanced.

Results on Hitachi SR2201 (U.Tokyo)

Parallel Performance of SB-BIC(0)-CG

2,471,439 DOF, 784,000 Elements, **$\lambda/E=10^6$**

Iterations/CPU time until convergence ($\varepsilon=10^{-8}$)

Precon-ditioning		16 PEs	32 PEs	48 PEs	64 PEs	96 PEs	144 PEs	192 PEs	256 PEs	Memory Size (GB)
BIC(0)	Iterations	14459	14583	15018	15321	15523	15820	16084	16267	
	sec.	13500	7170	4810	3630	2410	1630	1270	1230	3.10
	Speed-up	16	30	45	60	90	133	170	211	
BIC(1)	Iterations			379	390	402	424	428	452	
	sec.	N/A	N/A	236	175	119	81	62	48	8.39
	Speed-up			48	65	95	140	183	236	
BIC(2)	Iterations					364	387	398	419	
	sec.	N/A	N/A	N/A	N/A	212	140	112	86	14.4
	Speed-up					96	145	182	217	
SB-BIC(0)	Iterations	511	524	527	538	543	567	569	584	
	sec.	555	295	193	144	96	64	48	38	3.52
	Speed-up	16	30	46	62	92	139	185	235	

Results on Hitachi SR2201 (U.Tokyo)

Parallel Performance of SB-BIC(0)-CG

2,471,439 DOF, 784,000 Elements, **$\lambda/E=10^6$**

Iterations/CPU time until convergence ($\varepsilon=10^{-8}$)

Precon-ditioning		16 PEs	32 PEs	48 PEs	64 PEs	96 PEs	144 PEs	192 PEs	256 PEs	Memory Size (GB)
BIC(0)	Iterations	14459	14583	15018	15321	15523	15820	16084	16267	
	sec.	13500	7170	4810	3630	2410	1630	1270	1230	3.10
	Speed-up	16	30	45	60	90	133	170	211	
BIC(1)	Iterations			379	390	402	424	428	452	
	sec.	N/A	N/A	236	175	119	81	62	48	8.39
	Speed-up			48	65	95	140	183	236	
BIC(2)	Iterations					364	387	398	419	
	sec.	N/A	N/A	N/A	N/A	212	140	112	86	14.4
	Speed-up					96	145	182	217	
SB-BIC(0)	Iterations	511	524	527	538	543	567	569	584	
	sec.	555	295	193	144	96	64	48	38	3.52
	Speed-up	16	30	46	62	92	139	185	235	

Results on Hitachi SR2201 (U.Tokyo)

Parallel Performance of SB-BIC(0)-CG

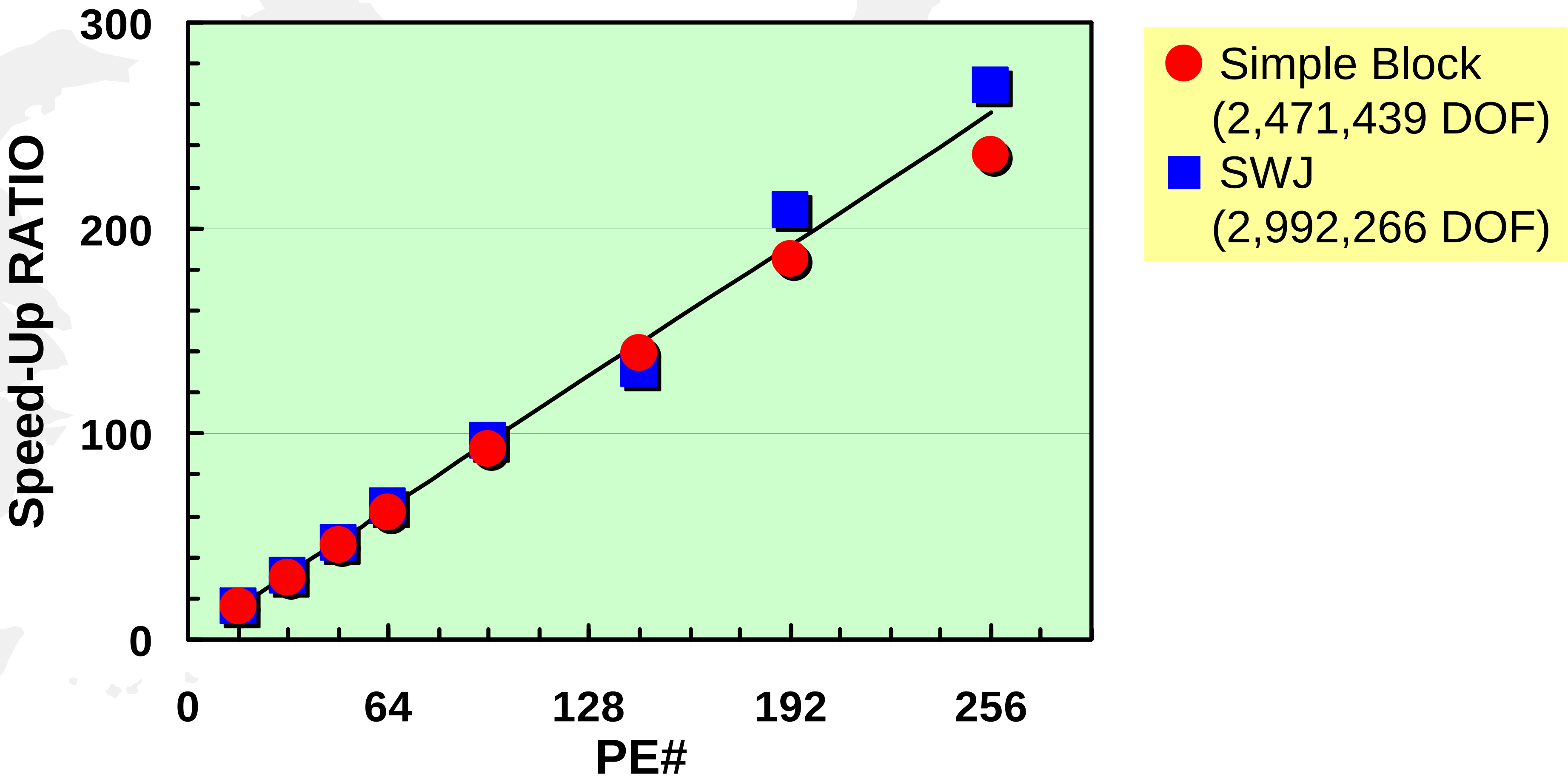
2,471,439 DOF, 784,000 Elements, **λ/E=10^6**

Iterations/CPU time until convergence ($\varepsilon=10^{-8}$)

Precon-ditioning		16 PEs	32 PEs	48 PEs	64 PEs	96 PEs	144 PEs	192 PEs	256 PEs	Memory Size (GB)
BIC(0)	Iterations	14459	14583	15018	15321	15523	15820	16084	16267	
	sec.	13500	7170	4810	3630	2410	1630	1270	1230	3.10
	Speed-up	16	30	45	60	90	133	170	211	
BIC(1)	Iterations			379	390	402	424	428	452	
	sec.	N/A	N/A	236	175	119	81	62	48	8.39
	Speed-up			48	65	95	140	183	236	
BIC(2)	Iterations					364	387	398	419	
	sec.	N/A	N/A	N/A	N/A	212	140	112	86	14.4
	Speed-up					96	145	182	217	
SB-BIC(0)	Iterations	511	524	527	538	543	567	569	584	
	sec.	555	295	193	144	96	64	48	38	3.52
	Speed-up	16	30	46	62	92	139	185	235	

Results on Hitachi SR2201 (U.Tokyo)

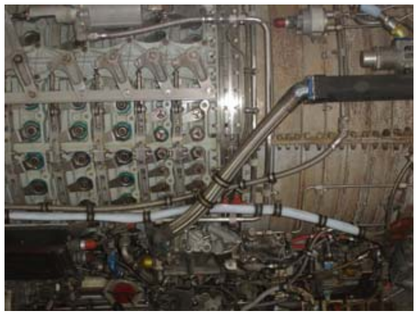
Parallel Performance of SB-BIC(0)-CG, λ/E=10^6

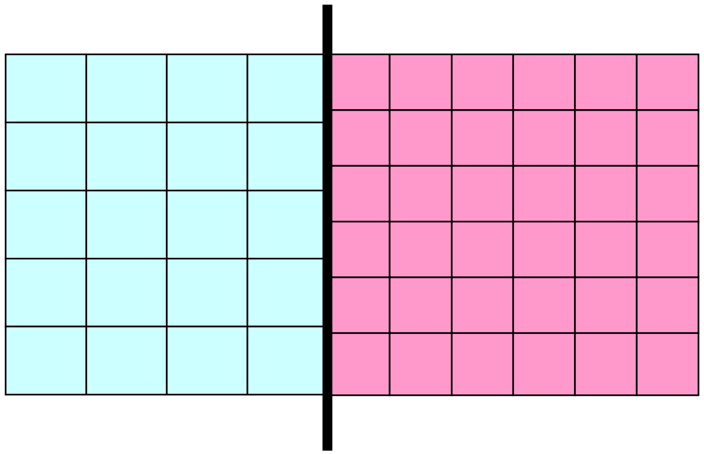


- Background
  - Simulations for Earthquake Generation Cycle
  - Selective Blocking
- **More General Problems**
  - **Extension of Overlapped Zones**
- Preconditioning/Partitioning Methods
  - Selective Fill-in
  - Selective Overlapping
- Results
- Summary
  - Future Works

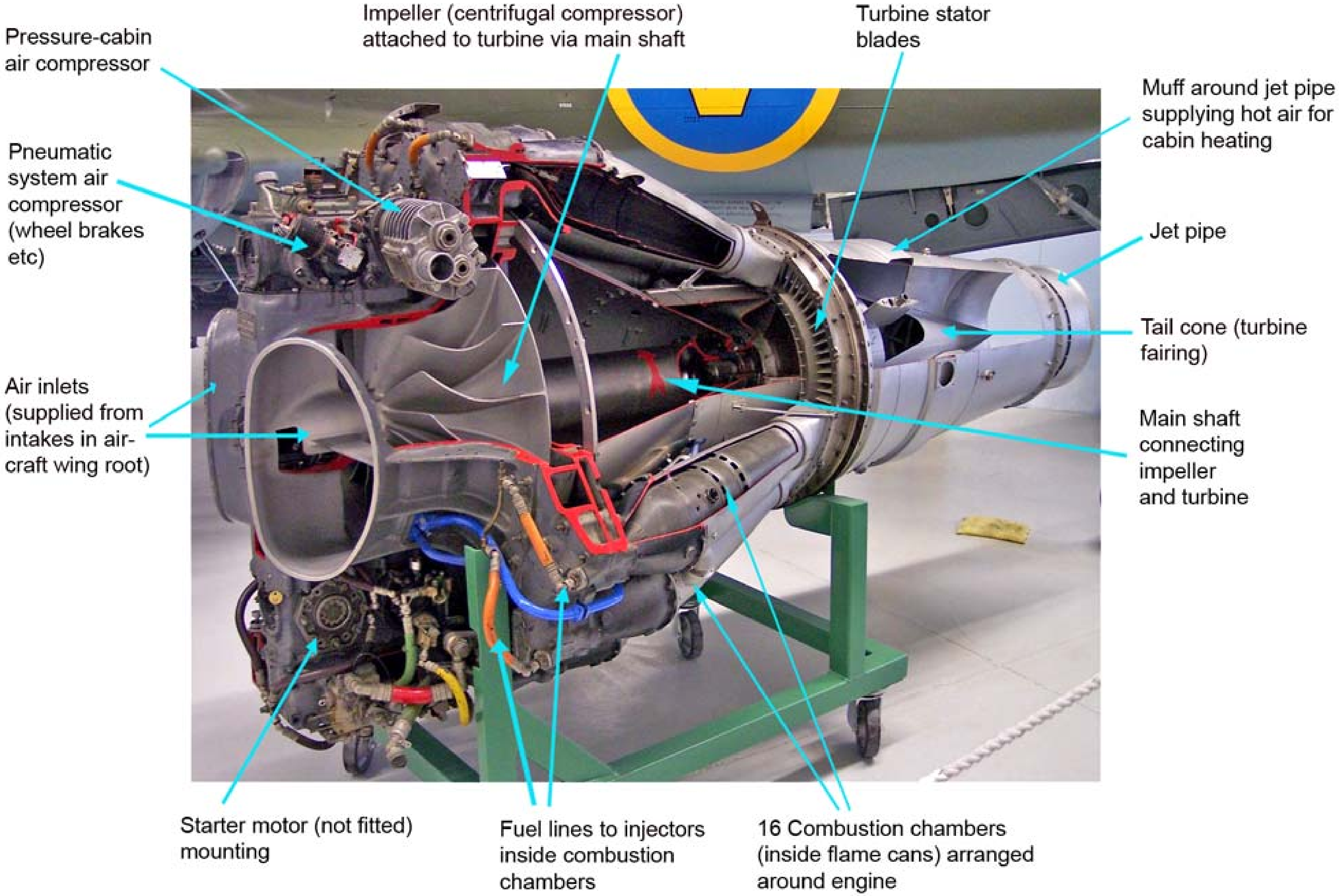
More General Problems

- Moving boundaries due to large slip conditions
- Inconsistent node number (and location) at boundary surfaces
  - Assembly structure for machine parts.
    - where meshes for each part are separately generated.
  - Commercial FEM codes (e.g. ABAQUS, NASTRAN) can treat problems for this type of "inconsistent" cases. (single PE, direct method for linear equations).





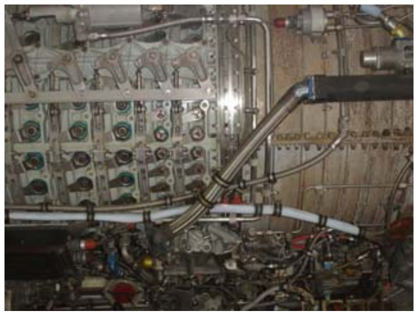
Example of Assembly Structure
Jet Engine

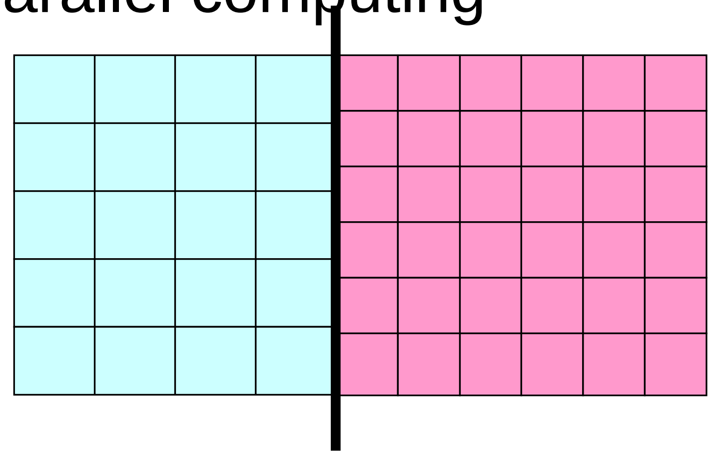


More General Problems

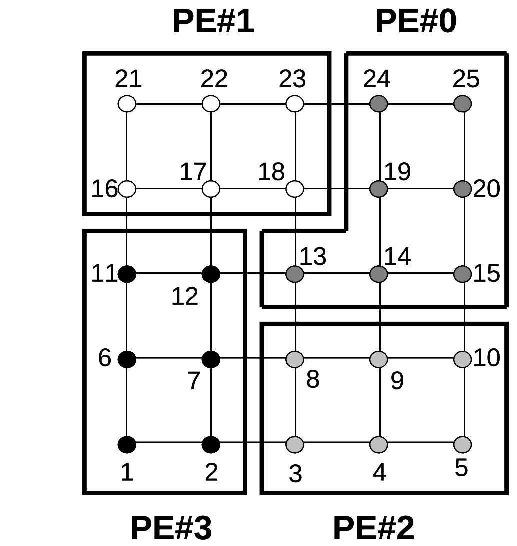
Inconsistent Number of Nodes at Boundary Surfaces

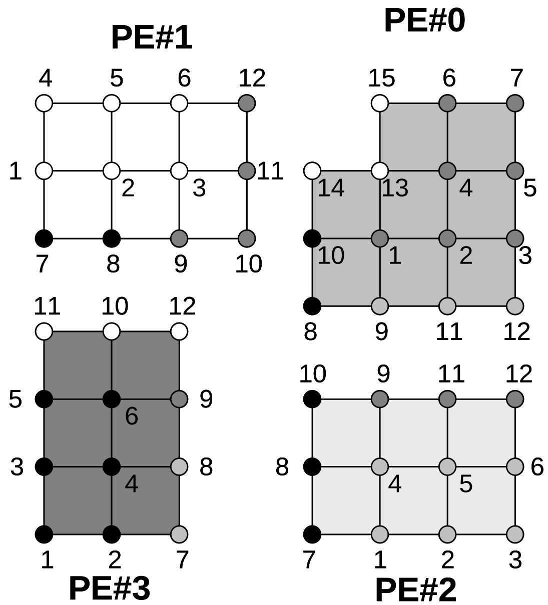
- Difficult to apply “selective blocking”
  - Size of each “selective block” may be too large for full LU factorization
- Difficult to apply “special partitioning”
- Remedy
  - Higher-order fill-in’s
  - Extension of overlapped zones for parallel computing



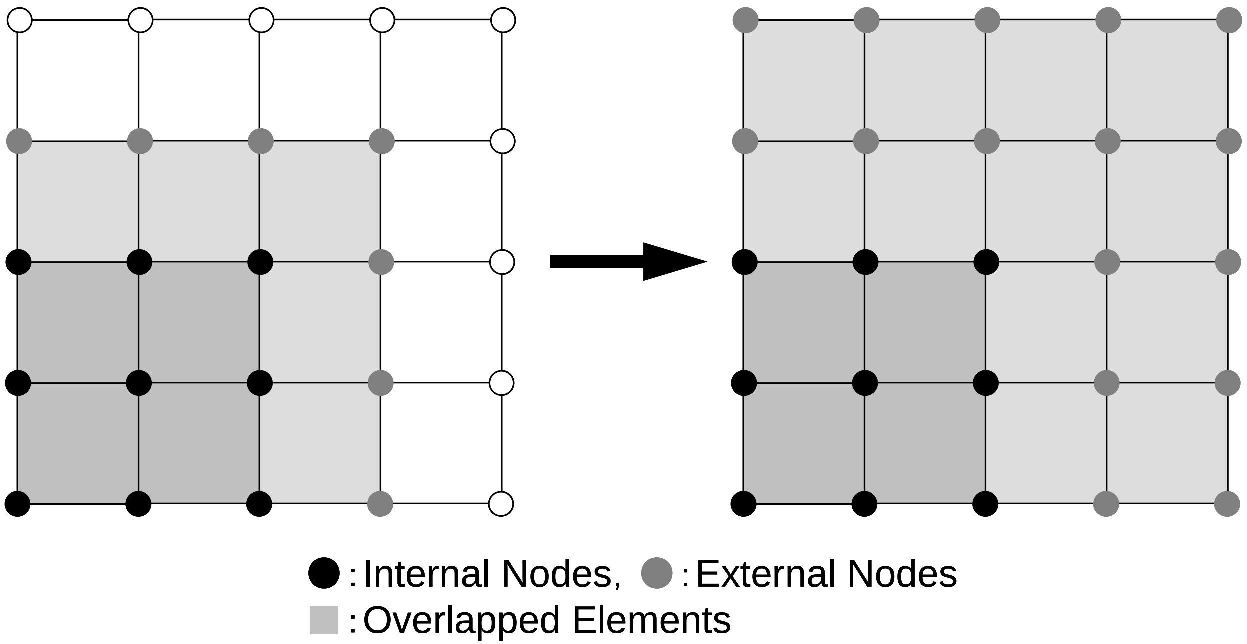


Extension of Overlapped Zones





Cost for computation and communication may increase



Effect of Extended Overlapped Zones

- [Nakajima, 2005]
  - BILU(0,1,2)
  - for “consistent” node number cases
  - IBM SP3 in NERSC/LBNL

Preconditioning	partitioning (overlap #)	PE #	iter’s	set-up+ solve(sec.)	parallel speed-up
SB-BILU (0)	special [3]	16	386	506.2	16.0
	1-layer	128	410	63.9	126.7
BILU(1)	special [3]	16	225	563.2	16.0
	1-layer	128	247	95.0	94.8
BILU(1)	regular	16	444	1033.2	16.0
	1-layers	128	529	191.0	86.6
BILU(1)	regular	16	405	1063.3	16.0
	2-layers	128	430	204.6	83.2
SPAI	regular	16	891	626.3	16.0
	2-layers	128	888	105.1	95.4

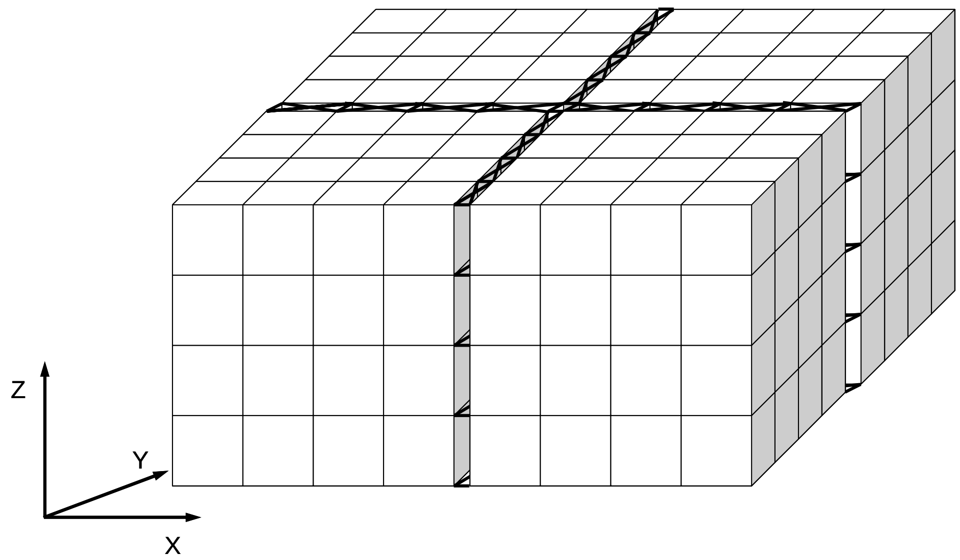
- Background
  - Simulations for Earthquake Generation Cycle
  - Selective Blocking
- More General Problems
  - Extension of Overlapped Zones
- **Preconditioning/Partitioning Methods**
  - **Target Application**
  - **Selective Fill-in**
  - **Selective Overlapping**
- Results
- Summary
  - Future Works

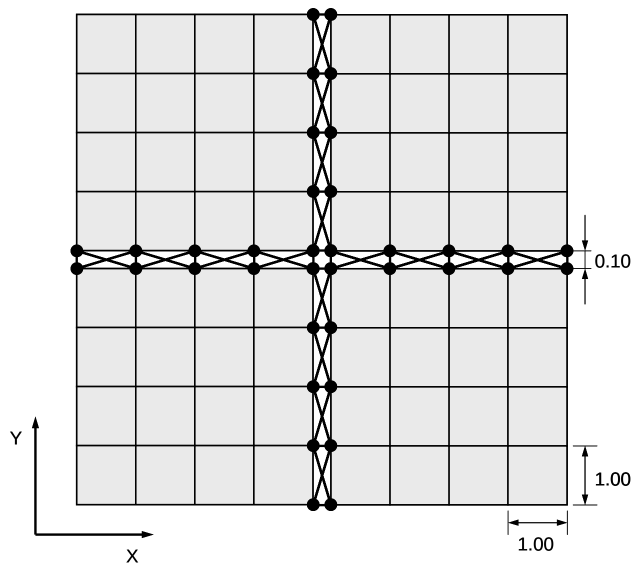
Robust and efficient preconditioning for parallel iterative solvers in more general cases

- Selective fill-in for serial & parallel computing
- Selective overlapping for parallel computing

Example for "Inconsistent" Cases

This model simulates contact problem in assembly structure

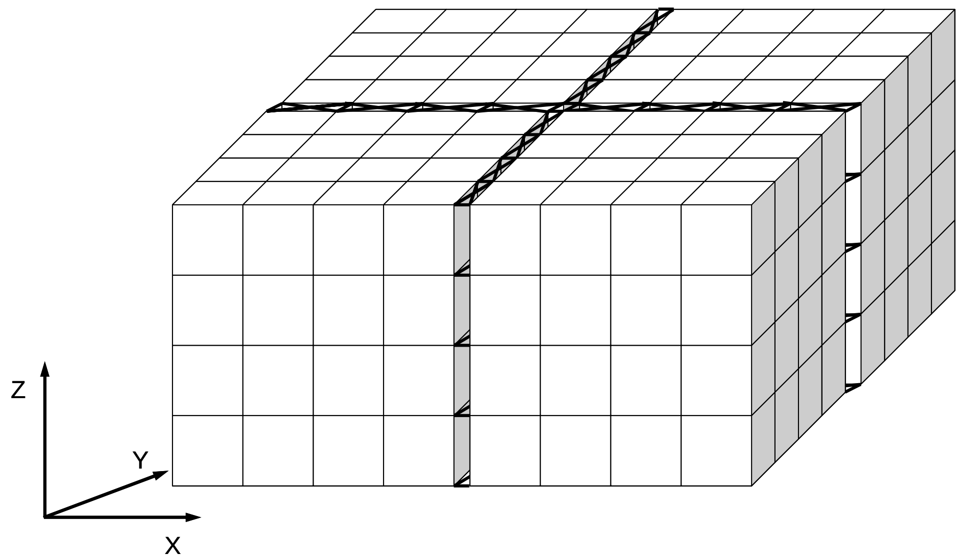


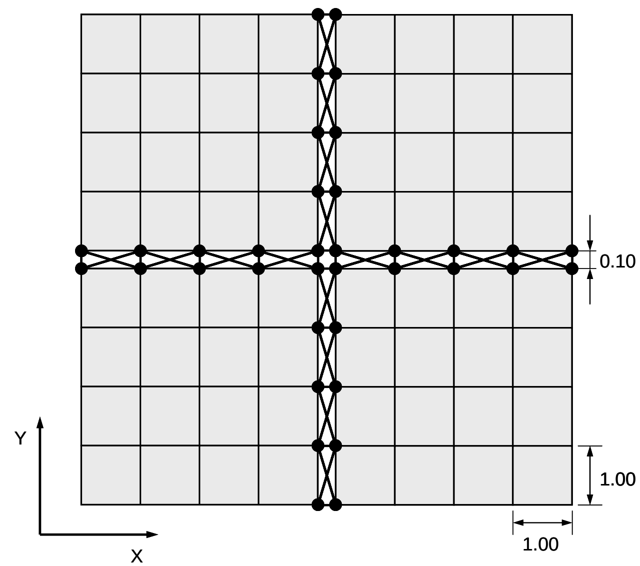


- Each block is discretized into cubic tri-linear elements
  - elastic material: E= 1.00, Poisson ration= 0.25
- Each block is connected through elastic truss elements generated on each node on contact surfaces.
  - Truss elements are crossing.

Example for "Inconsistent" Cases

This model simulates contact problem in assembly structure





- Elastic coefficient of truss elements is set to 10^3 times as large as that of solid elements.
  - This condition simulates constraint boundary conditions for contact.
- Distributed uniform force at $z=z_{max}$ surface
  - u=0@x=0, v=0@y=0, w=0@z=0

Selective Fill-in

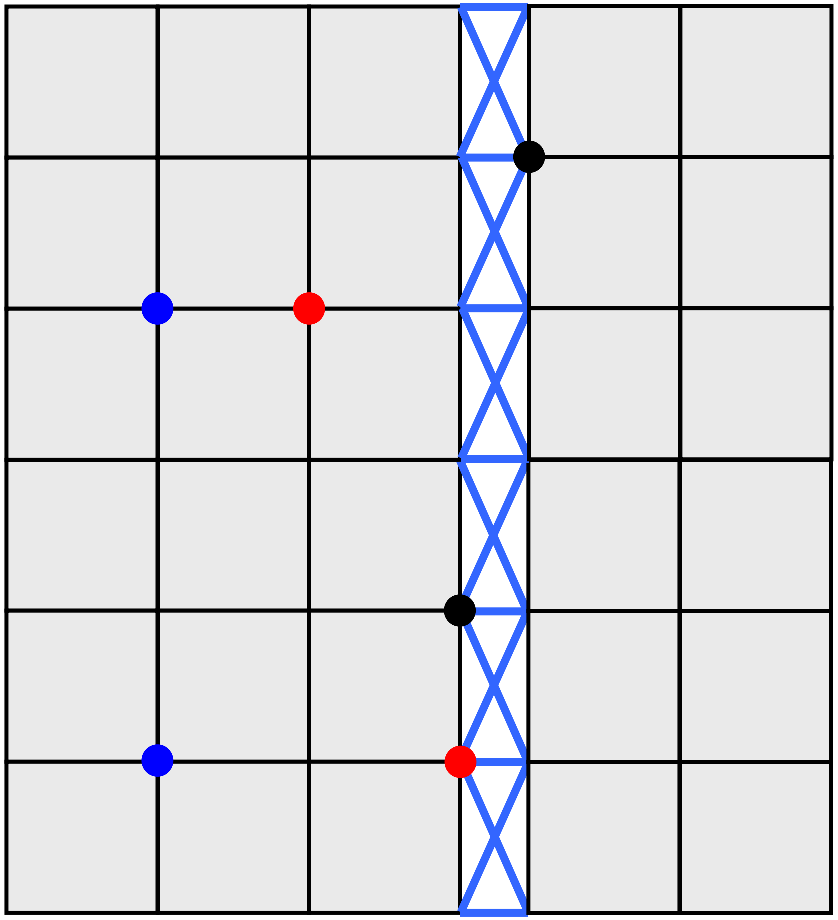
- Apply higher order of fill-in's between nodes which connect to truss-type elements.
  – Similar concept as "selective blocking"

- In this work: <u>**BILU(1+)**</u>
  – BILU(2) for these special nodes (2nd order fill-in's)
  – BILU(1) for general nodes (1st order fill-in's)
- Cost is similar to that of BILU(1), but effect of preconditioning is expected to be competitive with that of BILU(2).

What is “fill-in” ?

- Coefficient matrices [A] of [A]{x}={b} for finite-element applications are generally sparse.
- But inverse matrices of [A] are not necessarily sparse.
  - “Fill-in” occurs.
- If all fill-in’s are allowed.
  - Full LU factorization, Gaussian Elimination
  - Robust but expensive
- If no fill-in’s are allowed.
  - Same sparsity as [A]
  - Incomplete LU factorization with 0-level fill-in’s = ILU(0)
  - ILU(1), ILU(2) …

Idea of “Selective Fill-in”: ILU(1+)



- 2nd order fill-in’s are considered for these nodes (black)
- 2nd order fill-in’s are NOT considered for these nodes (red)
- 2nd order fill-in’s are NOT considered for these nodes (blue)

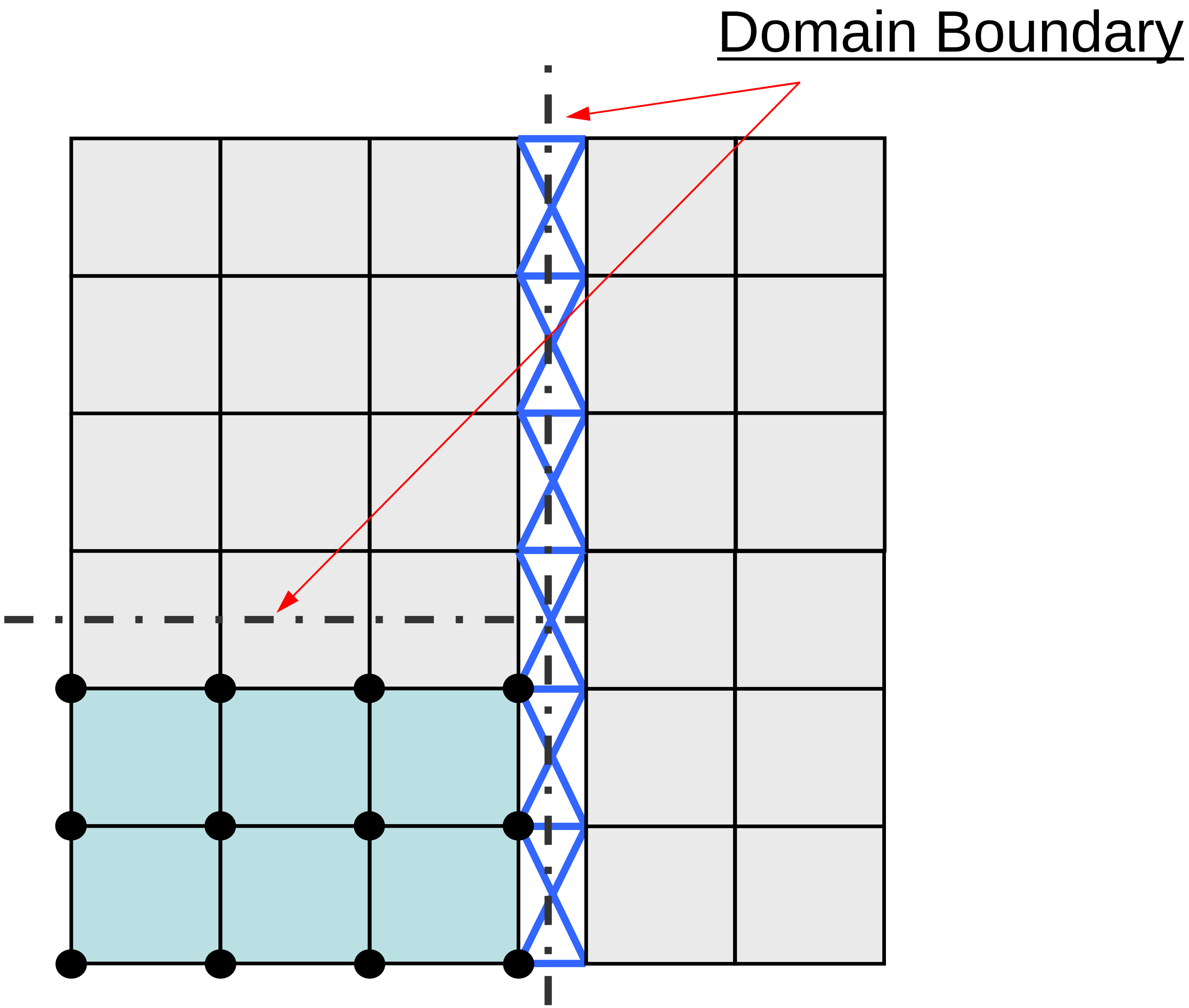
Selective Overlapping

- Same rules in “selective fill-ins” are applied to extention of overlapping zones.
  - Similar concept as “selective blocking”

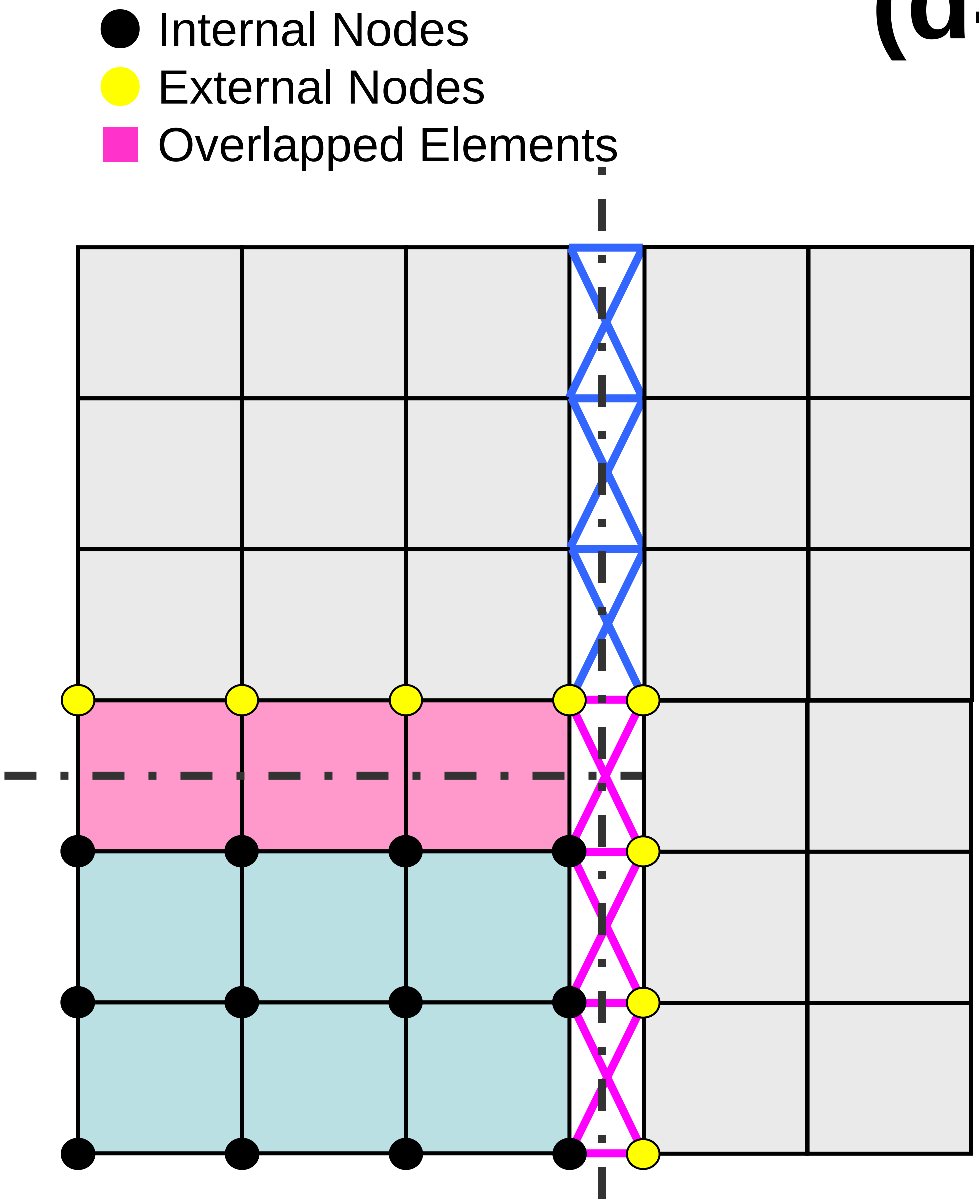
- In selective overlapping, extension of overlapping for nodes that are not connected to special elements for contact conditions is *delayed*.
- The increase in cost for computation and communication by extension of overlapped elements is suppressed.

Internal Nodes for Partitioning

● Internal Nodes

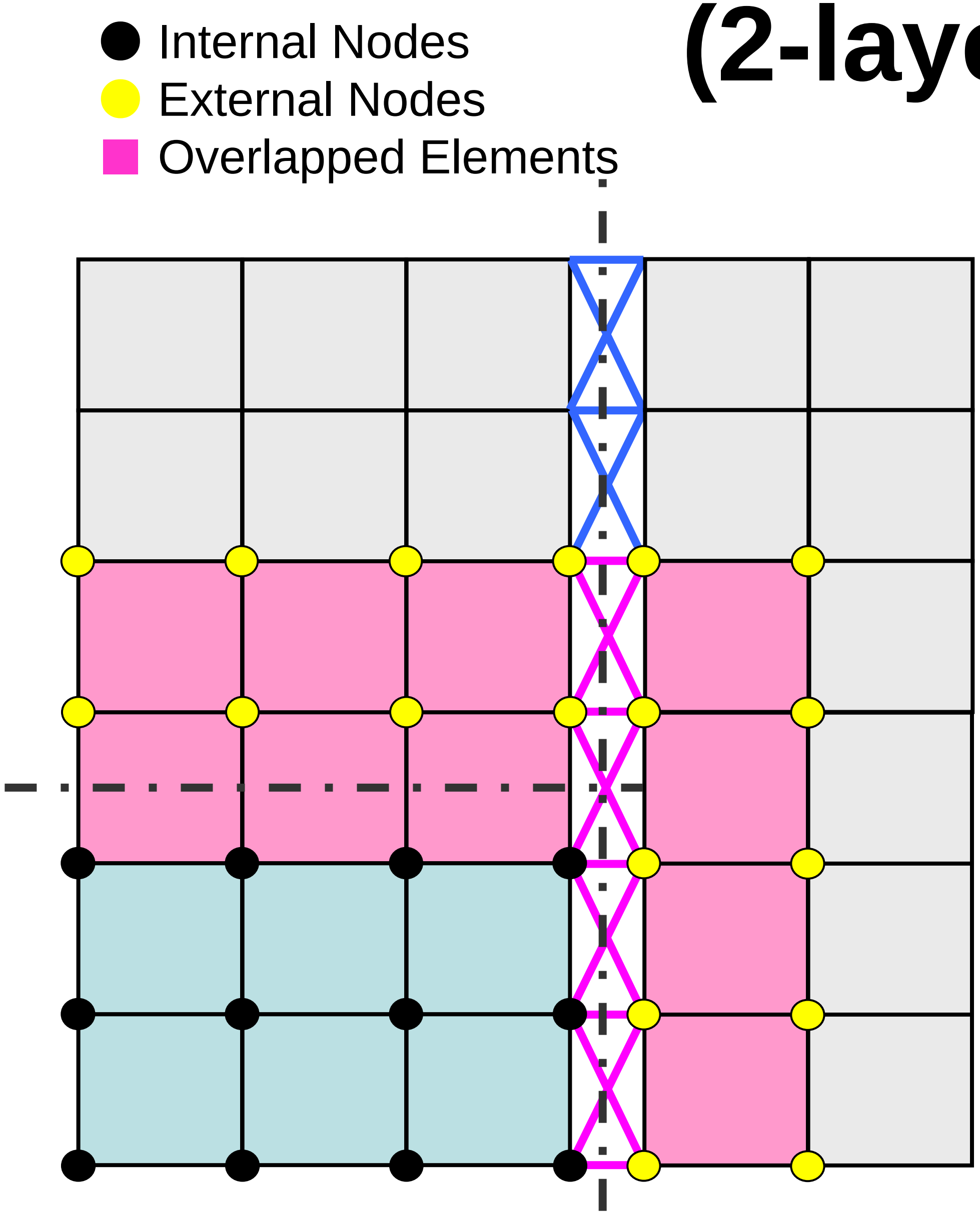


One-Layer Overlapping (d=0/1)



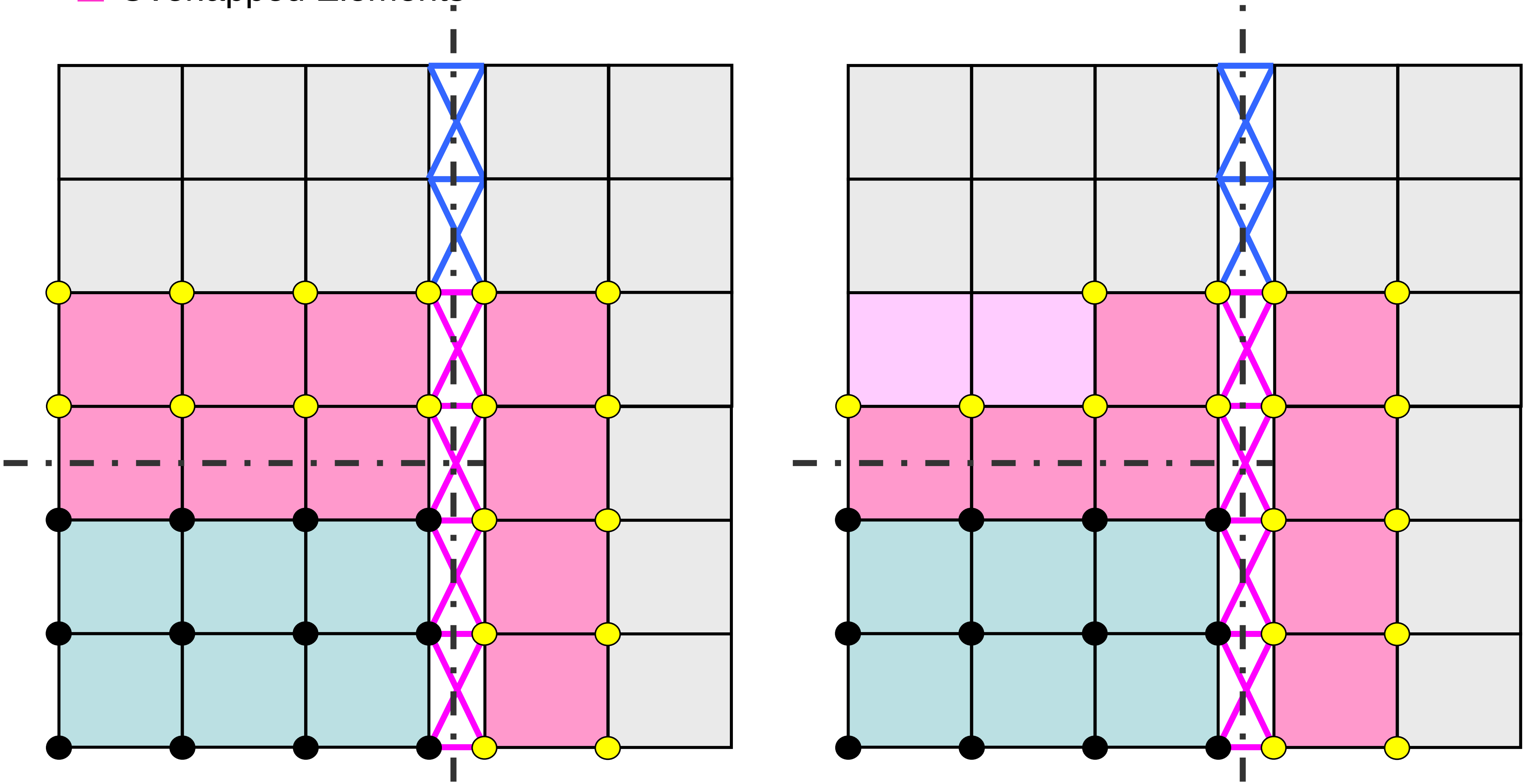
This is the general configuration of local data set for parallel FEM (one-layer of overlapping).

Extension of Overlapped Zones (2-layers: d=2)



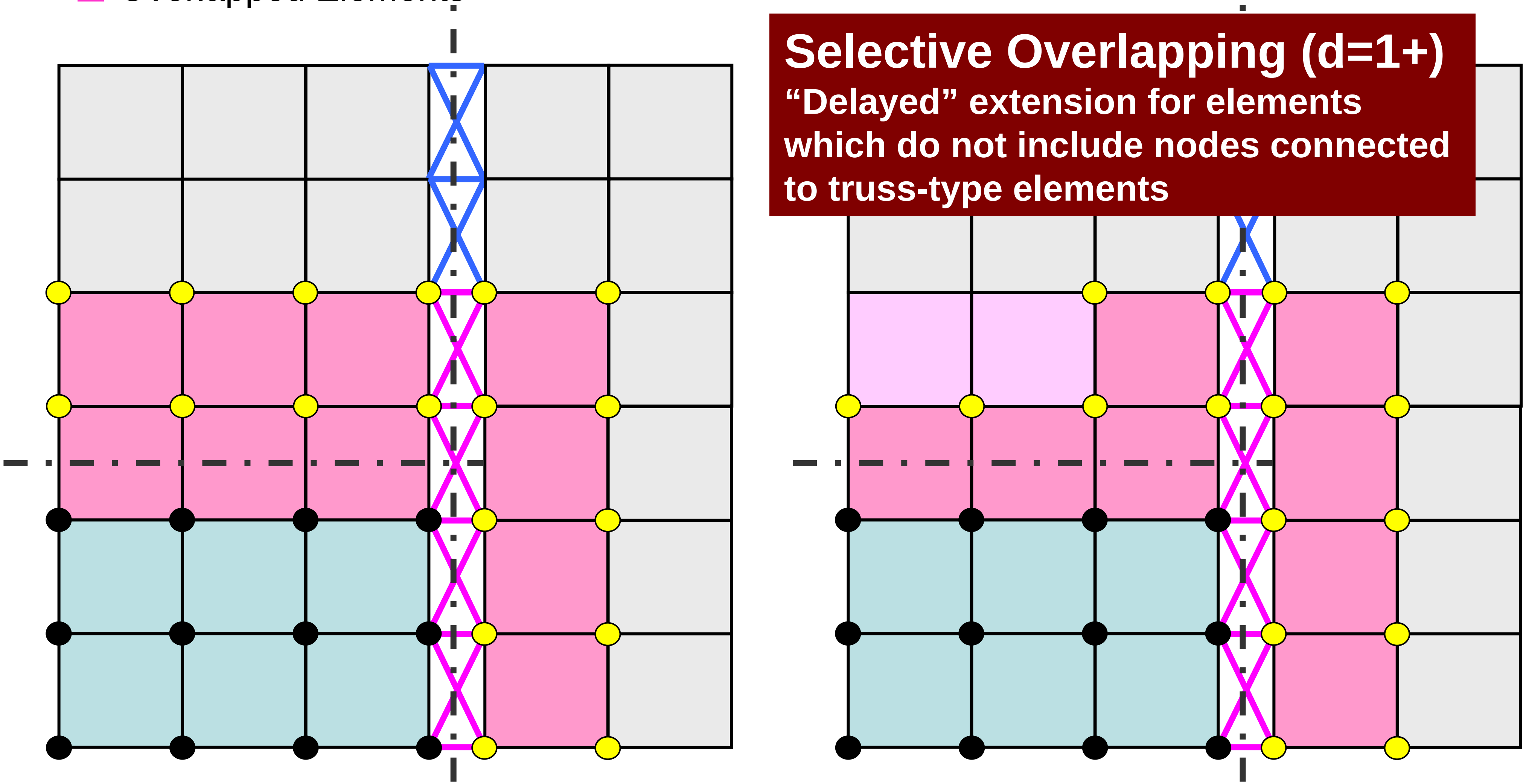
Extension of Overlapped Zones (d=2 and d=1+)

- Internal Nodes
- External Nodes
- Overlapped Elements



Extension of Overlapped Zones (d=2 and d=1+)

● Internal Nodes
● External Nodes
■ Overlapped Elements



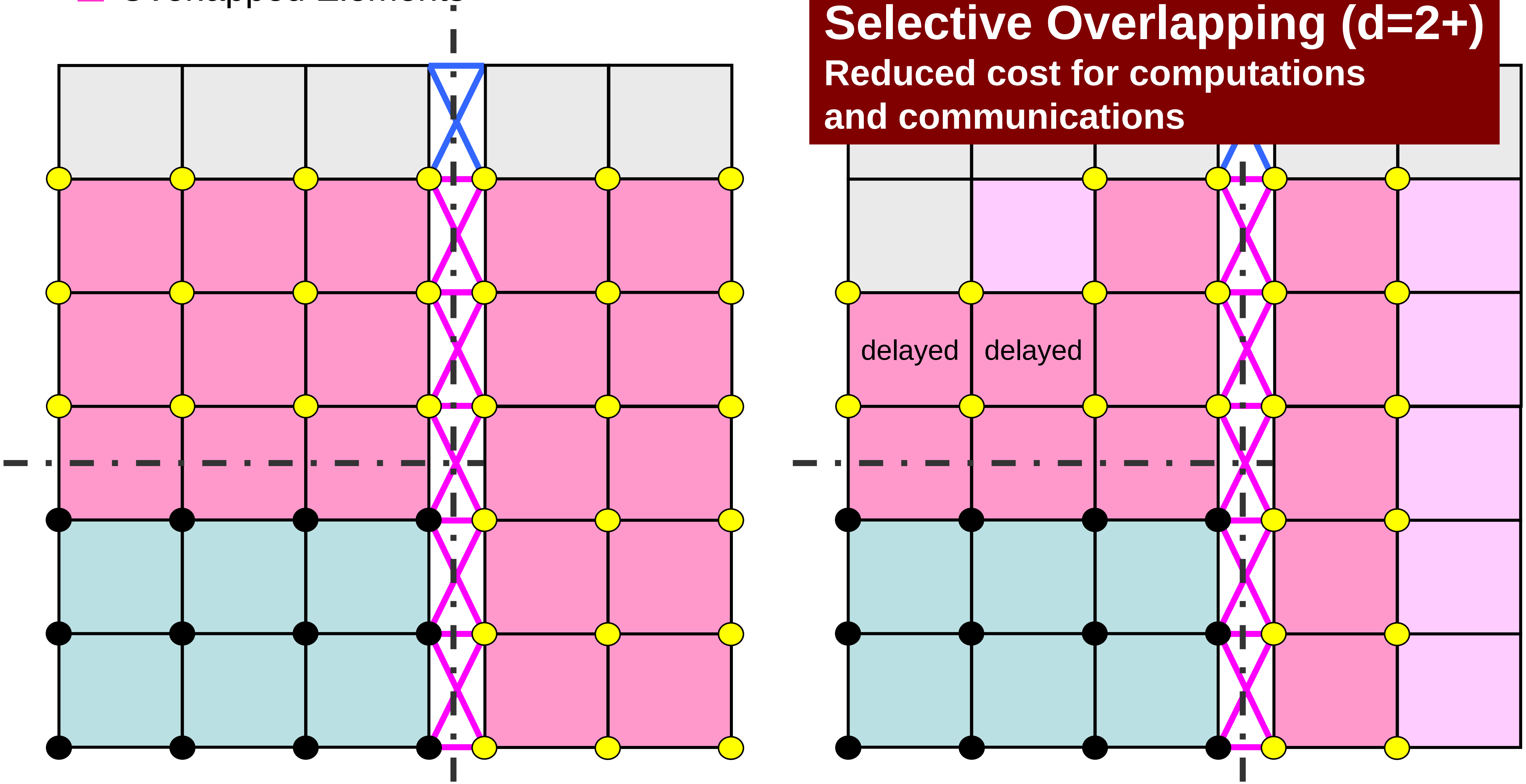
Selective Overlapping (d=1+)
"Delayed" extension for elements which do not include nodes connected to truss-type elements

Extension of Overlapped Zones (d=3 and d=2+)

- ● Internal Nodes
- ● External Nodes
- ■ Overlapped Elements

Selective Overlapping (d=2+)

Reduced cost for computations and communications



BILU with selective fill-in/overlapping

- **BILU (p)-(d)**
  - **p** level of fill-ins (0, 1, 1+, 2, 2+ …)
  - **d** depth of overlapping (0, 1, 1+, 2, 2+ …)

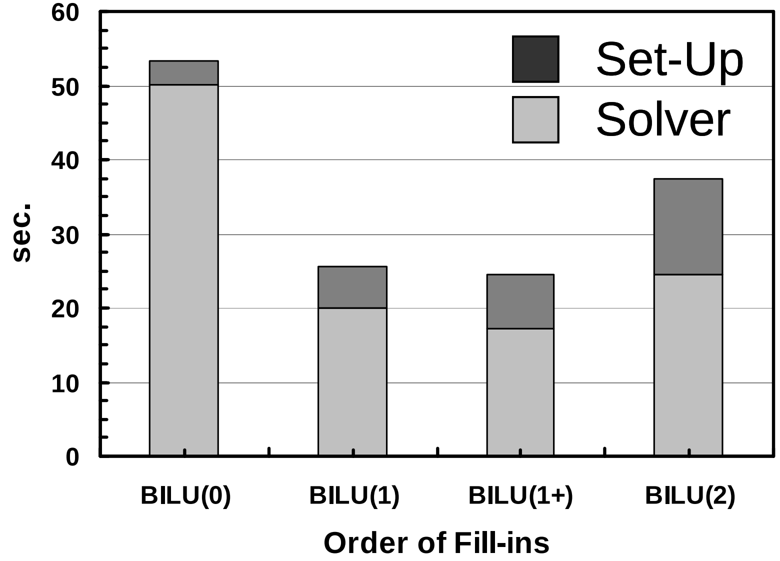
- Background
  - Simulations for Earthquake Generation Cycle
  - Selective Blocking
- More General Problems
  - Extension of Overlapped Zones
- Preconditioning/Partitioning Methods
  - Target Application
  - Selective Fill-in
  - Selective Overlapping
- **Results**
- Summary
  - Future Works

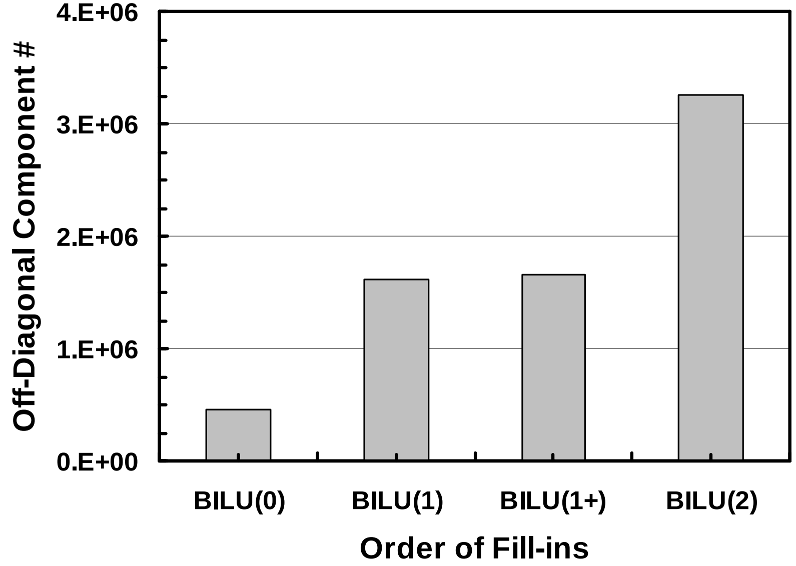
Summary of Problem Setting

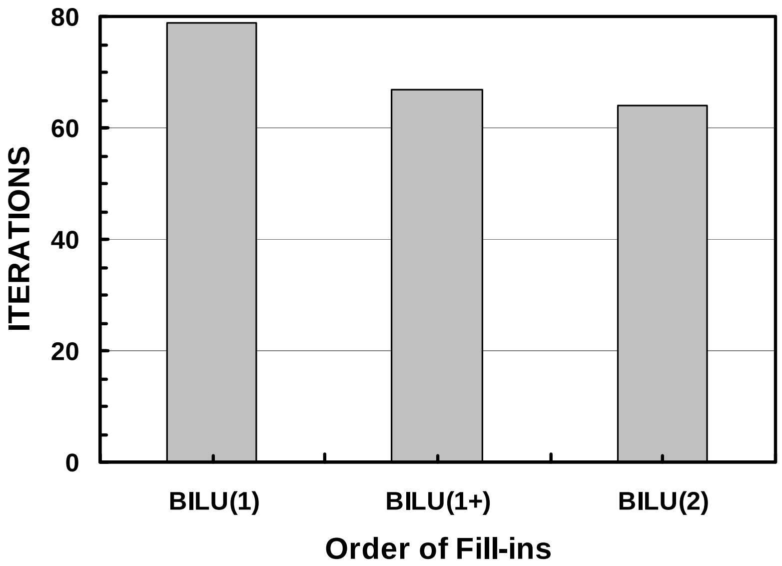
- Problem Size
  - Small: 38,148 elements (except truss's) 117,708 DOF
  - Large: 1,000,000 elements (except truss's), 3,152,412 DOF
- Preconditioned GPBiCG [Zhang, 1997]
  - for general matrices, although the matrices are SPD
  - Localized preconditioning (block Jacobi type)
    - BILU（0,1,2）, **Selective Fill-in (BILU(1+))**
- Partitioning
  - Recursive Coordinate Bisection（RCB）: 8~64
    - **Selective Overlapping**
- Environment
  - 64-core AMD Opteron 275 (2.2GHz), Infiniband
  - F90 + MPI

Results: Small Case (1PE)

117,708 DOF, $\lambda=10^3$, $\varepsilon=10^{-8}$



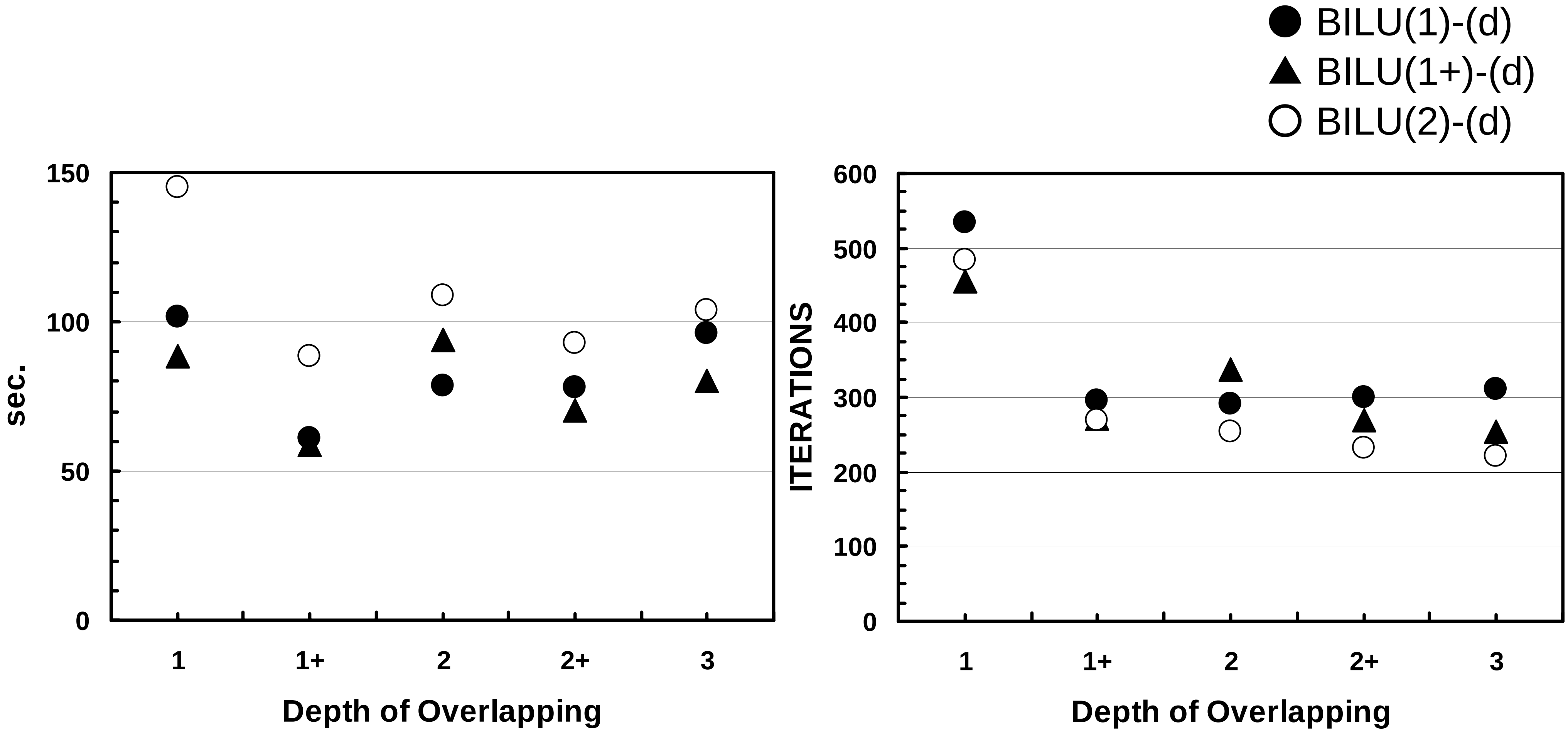




320 iterations
for BILU(0)

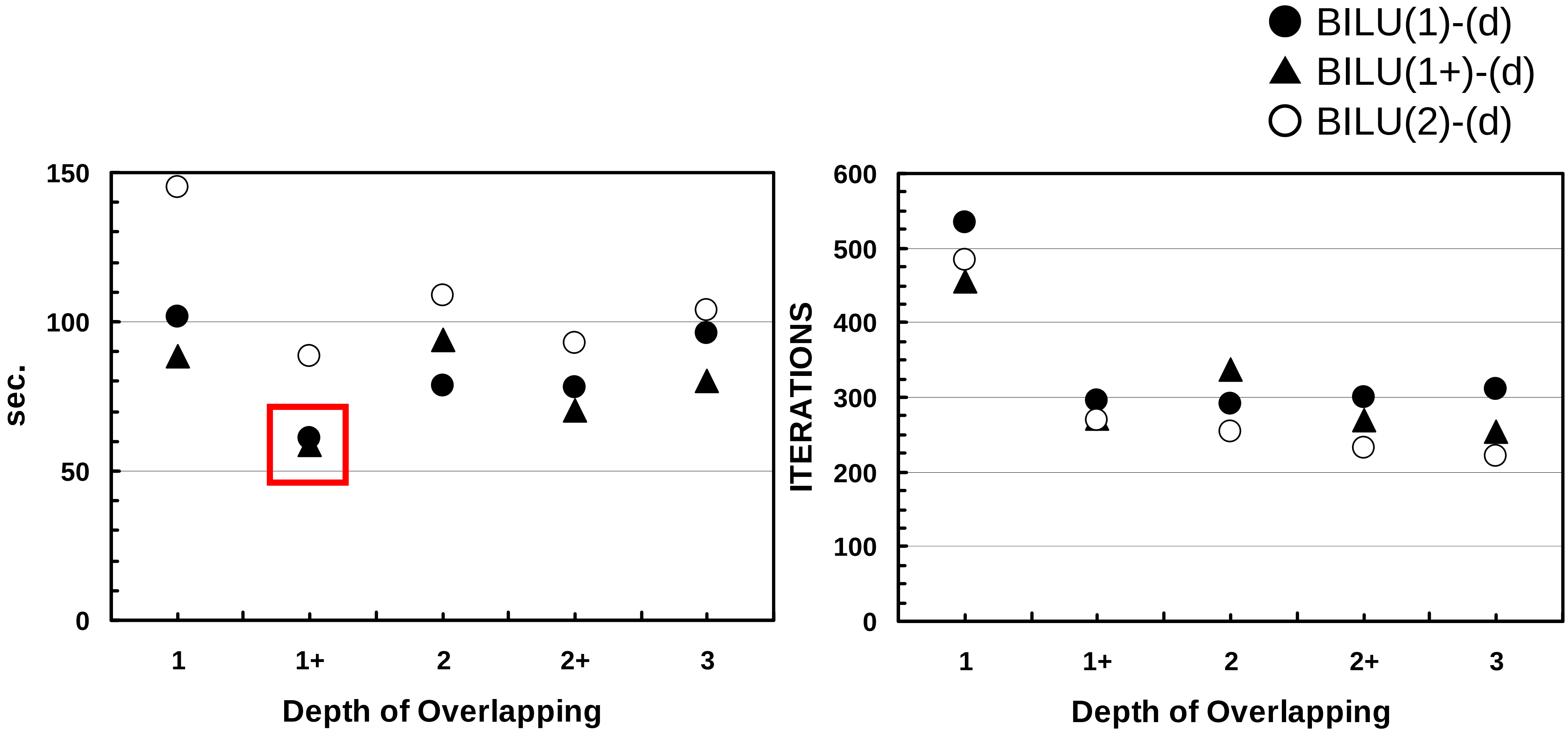
Results: Large Cases (64 cores)

3,152,412 DOF , $\lambda=10^3$, $\varepsilon=10^{-8}$



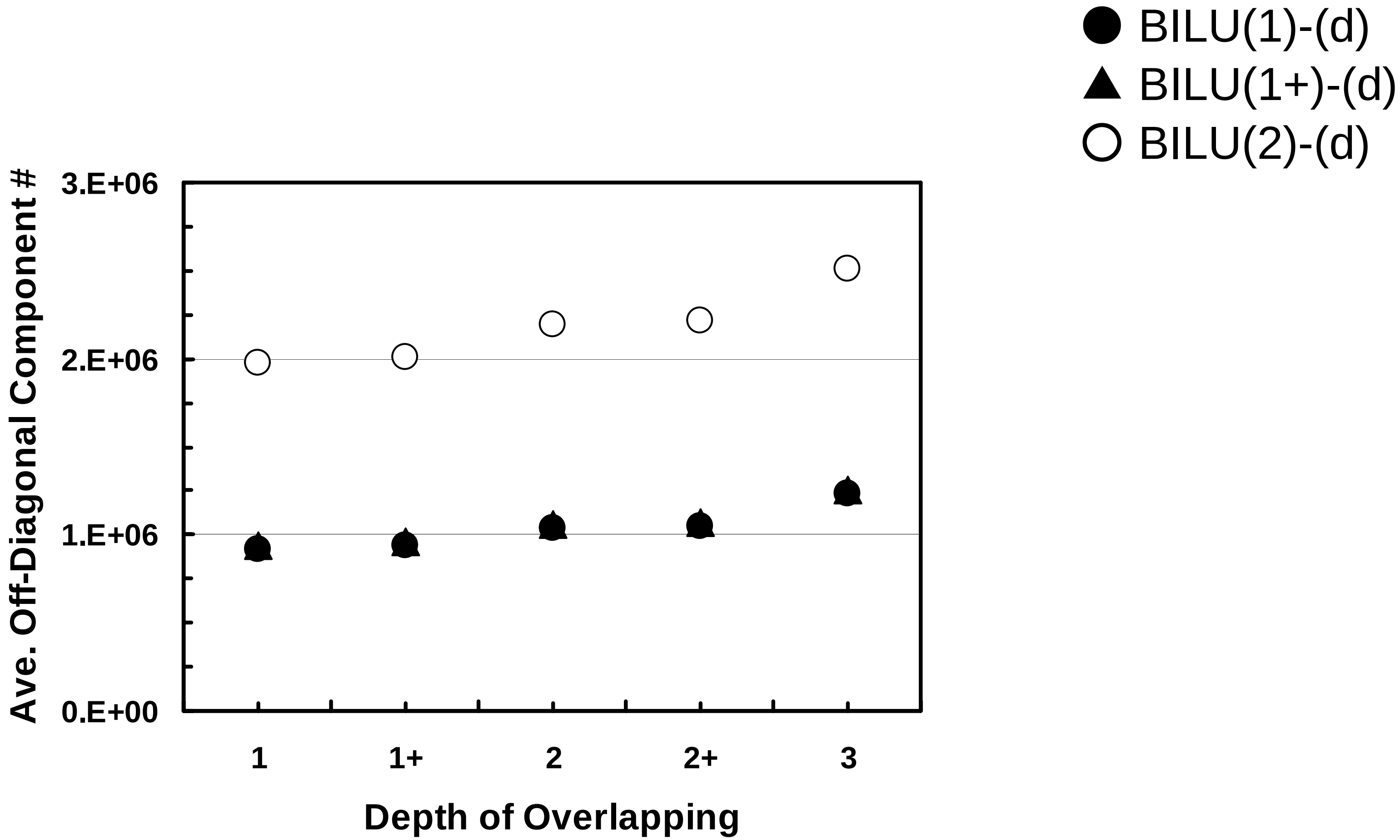
Results: Large Cases (64 cores)

3,152,412 DOF , $\lambda=10^3$, $\varepsilon=10^{-8}$



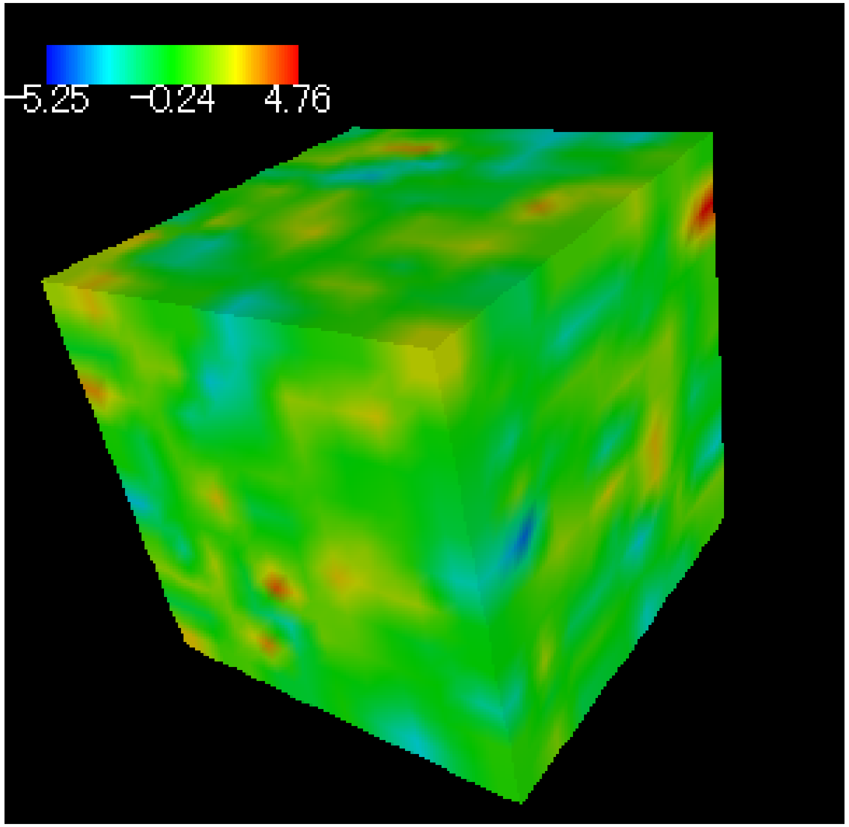
Results: Large Cases (64 cores)

3,152,412 DOF , $\lambda=10^3$, $\varepsilon=10^{-8}$



Elastic Cube with Heterogeneous Distribution of Material Property

- E= 1.e-2 ~ 1.e+2, ν= 0.25
- Selective Fill-in/Overlapping with Threshold (for E)
  - **BILU (p,ω)-(d,α)**
- Scalability
  - Strong Scaling
    - 3,090,903 DOF, 1,000,000 cubes
  - Weak Scaling
- Hardware
  - Opeteon Cluster
  - TSUBAME Grid Cluster
    - up to 512 cores



TSUBAME Grid Cluster

Tokyo Institute of Technology

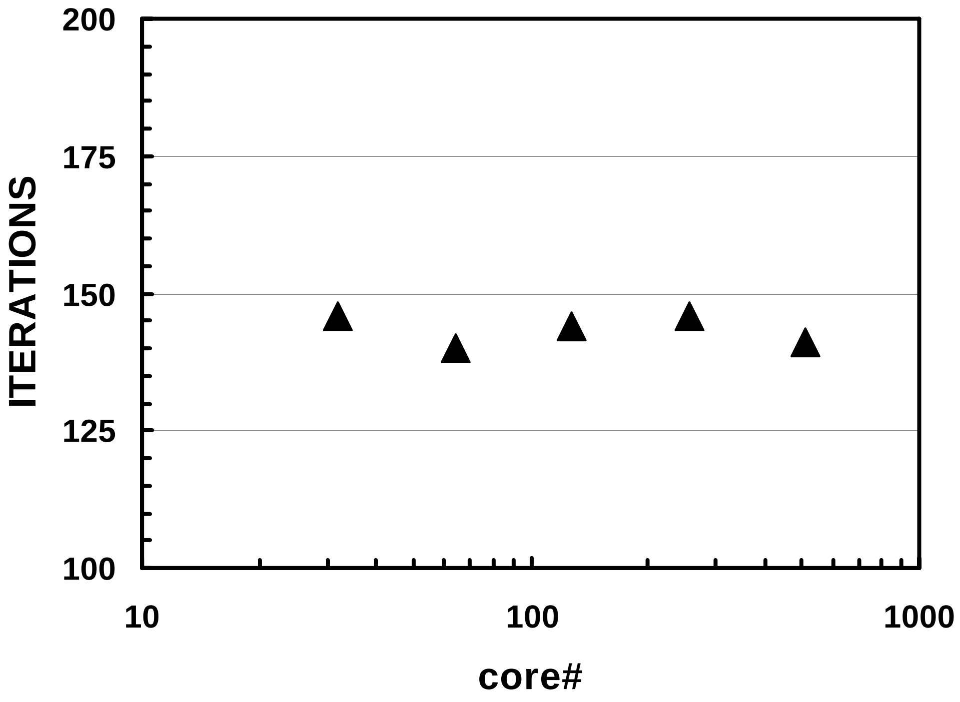
http://www.gsic.titech.ac.jp/

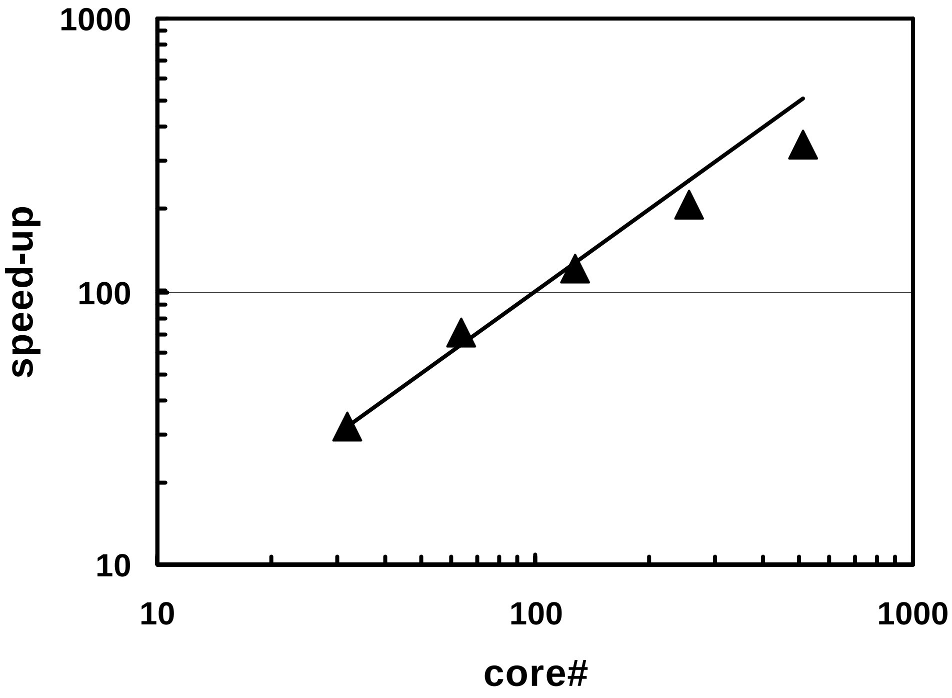
- A scalar system with 655 SunFire X4600 nodes, where each node has 8 sockets (16 cores) of AMD's dual-core-Opteron at 2.4 GHz, connected through Coherent HyperTransport.
  - Entire system has 10,480 cores, and 21.4 TB memory.
- Total peak performance is 50.4 TFLOPS.
  - 9th in 2006-NOV TOP500 list (47.38 TFLOPS with ClearSpeed Accelerators) (1st place in Japan)
- SMP nodes are connected through Infiniband 4x/Voltaire ISR 9288 switch.
- In the present work, only 8 cores on each SMP node have been used

Heterogeneous Elastic Cube

3,090,903 DOF, Strong Scaling on TSUBAME Grid Cluster (up to 512 cores)

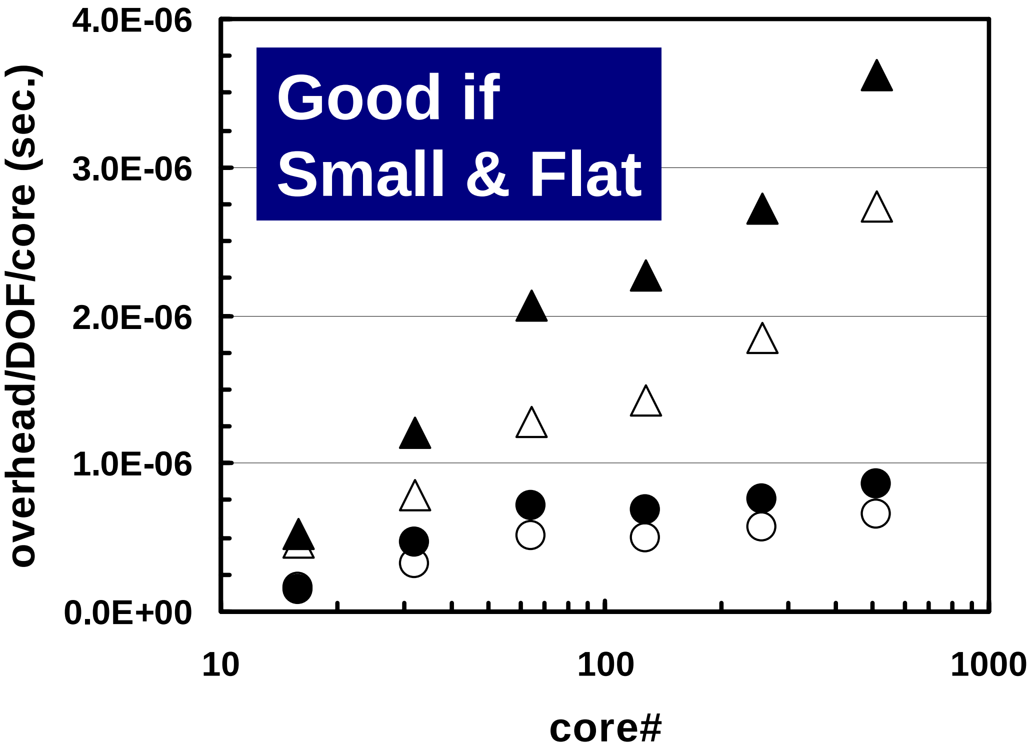
BILU(p,ω)-(d,α) ⇒ BILU(1+,5.)-(1+, 10.)





Effect of Selective Overlapping on Performance (Comm. Overhead)

Weak Scaling on TSUBAME Grid Cluster (up to 512 cores)



98,304 DOF/core

○ BILU(1+,5)-(1)

● BILU(1+,5)-(1+,10)

6,591 DOF/core

△ BILU(1+,5)-(1)

▲ BILU(1+,5)-(1+,10)

If the problem size per each core is sufficiently large, the additional overhead by *selective overlapping* (**(d=1)** and **(d+,α)= (1+,10)**) is negligible.

- Background
  - Simulations for Earthquake Generation Cycle
  - Selective Blocking
- More General Problems
  - Extension of Overlapped Zones
- Preconditioning/Partitioning Methods
  - Target Application
  - Selective Fill-in
  - Selective Overlapping
- Results
- **Summary**
  - **Future Works**

Summary

- Preconditioning method for ill-conditioned problems has been developed.
  - “Selective Fill-in”
  - “Selective Overlapping”

- Developed method provides robust and efficient convergence with excellent scalability for ill-conditioned problems:
  - contact problem
  - heterogeneous material property

Future Works

Towards more robust parallel preconditioning …

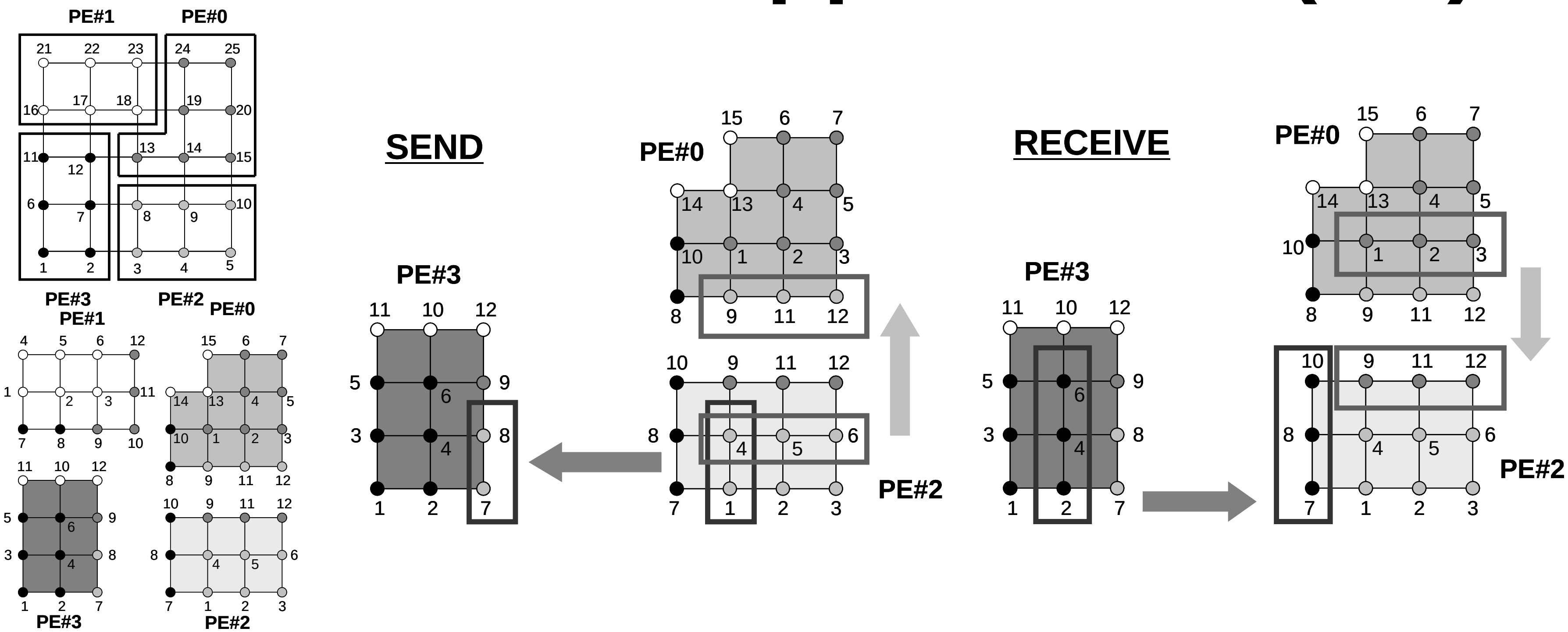
- Global reordering method for finding independent sets in distributed data sets is to be developed.
- Some hierarchical approach should be useful.
  - PHIDAL [Henon & Saad, 2007]

- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- Overview of the Current Project by JST
  - Integrated Predictive Simulation System for Earthquake and Tsunami Disaster
- **Some Technical Issues**
  - Parallel Preconditioning Methods
  - **Vector vs. Scalar Processors**
  - **Parallel Programming Models in Multi-Core Era**
- Future Directions

Features of FEM applications (1/2)

- Most of the computation time is spent for matrix assembling/formation and solving linear equations.
- **HUGE** “indirect” accesses
  - memory intensive
  - element-by-element
  - vertex-by-vertex (node-by-node)
- Local “element-by-element” operations
  - sparse coefficient matrices
  - suitable for parallel computing

Features of FEM applications (2/2)



- In parallel computation …
  - communications with ONLY neighbors (except “dot products” etc.)
  - amount of messages are relatively small because only values on domain-boundary are exchanged.
  - communication (MPI) latency is critical

Optimization of ICCG Solvers

- IC (Incomplete Cholesky)/ILU (Incomplete LU) factorization are general and robust preconditioning method for iterative linear solvers.
- Forward/backward substitution processes in IC/ILU are not suitable for parallel/vector computers. But it's possible with:
  - block-Jacobi type localized preconditioning for parallel computation with MPI.
  - **multicolor ordering**
  - **proper method for storage of matrix coefficients.**
    - **DJDS**
    - **DCRS**

$$y_k = b_k - \sum_{j=1}^{k-1} l_{kj} y_j \ (k = 2, \cdots, N)$$

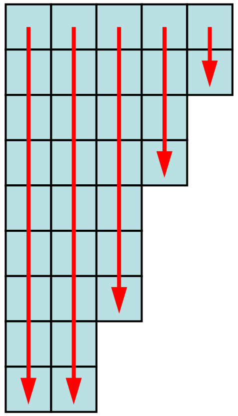
$$x_k = \tilde{d}_k \left(y_k - \sum_{j=k+1}^{N} u_{kj} y_j \right) \quad (k = N, N-1, \cdots, 1)$$

Matrix Storage/Structure of Loops

- DJDS (Descending order Jagged Diagonal Storage) with long innermost loops is suitable for vector processors.

- Reduction type loop of DCRS is more suitable for cache-based scalar processor because of its localized operation.

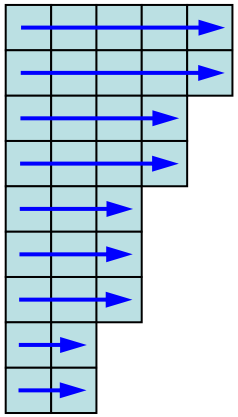
DJDS



```
do iv= 1, NVECT
  iv0= STACKmc(iv-1)
  do j= 1, NLhyp(iv)
    iS= index_L(NL*(iv-1)+ j-1)
    iE= index_L(NL*(iv-1)+ j  )
    do i= iv0+1, iv0+iE-iS
       k= i+iS - iv0
      kk= item_L(k)
      Z(i)= Z(i) - AL(k)*Z(kk)
    enddo

    iS= STACKmc(iv-1) + 1
    iE= STACKmc(iv  )
    do i= iS, iE
      Z(i)= Z(i)/DD(i)
    enddo
  enddo
enddo
```

DCRS

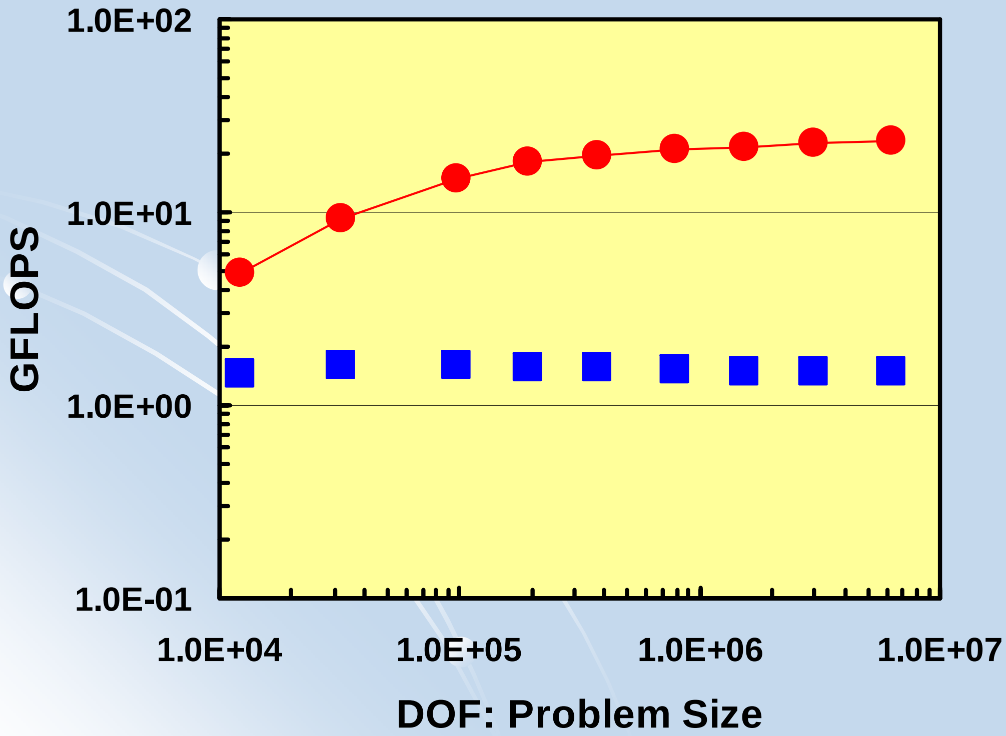


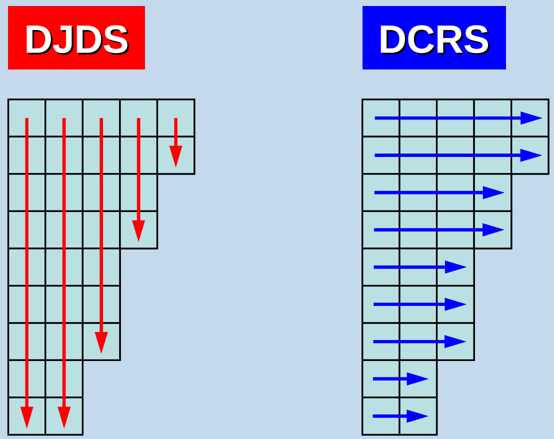
```
do i= 1, N
  SW= WW(i,Z)
  isL= index_L(i-1)+1
  ieL= index_L(i)
  do j= isL, ieL
    k = item_L(j)
    SW= SW - AL(j)*Z(k)
  enddo

  Z(i)= SW/DD(i)
enddo
```

Impact of Reordering on 3D Elastic Simulation (FEM) Problem Size~GFLOPS

Earth Simulator, 8 PE's, Flat-MPI

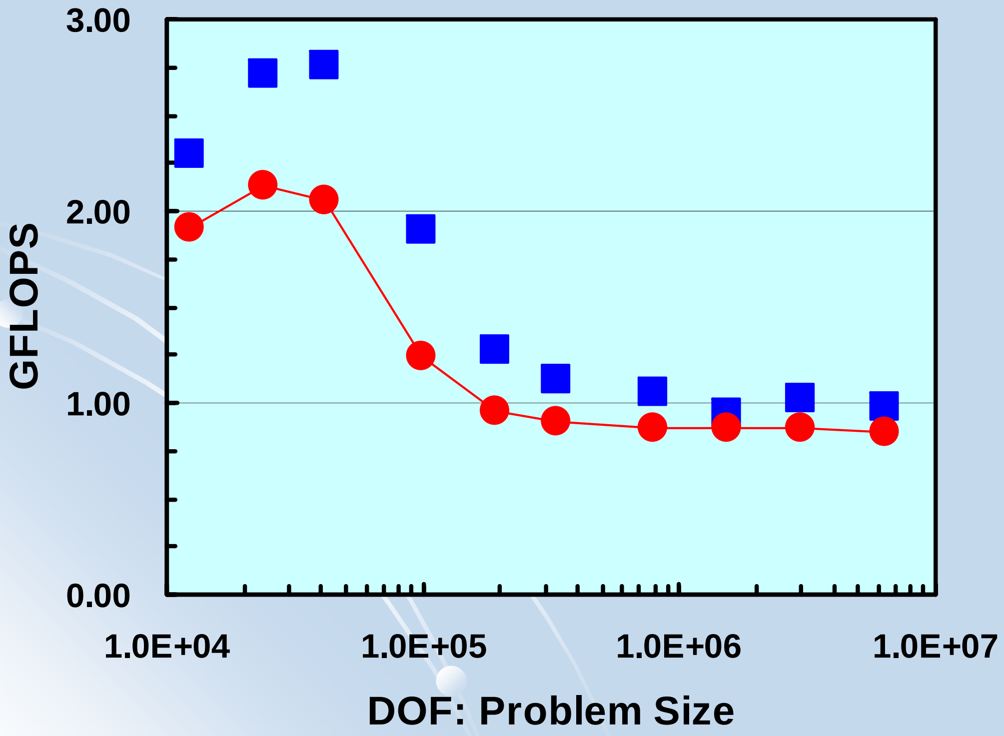


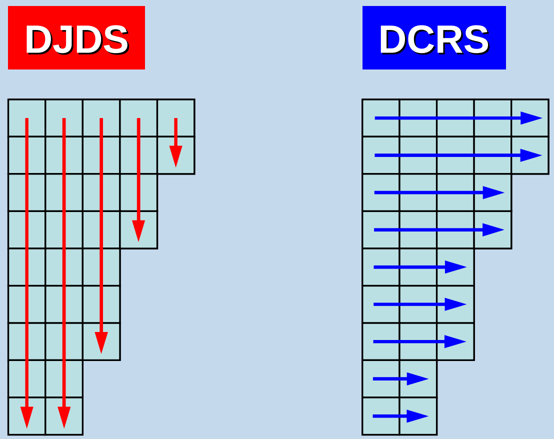


DJDS (Descending order Jagged Diagonal Storage) with long innermost loops is suitable for vector processors.

Impact of Reordering on 3D Elastic Simulation (FEM) Problem Size~GFLOPS

IBM SP-3 (Seaborg@NERSC), 8 PE's, Flat-MPI

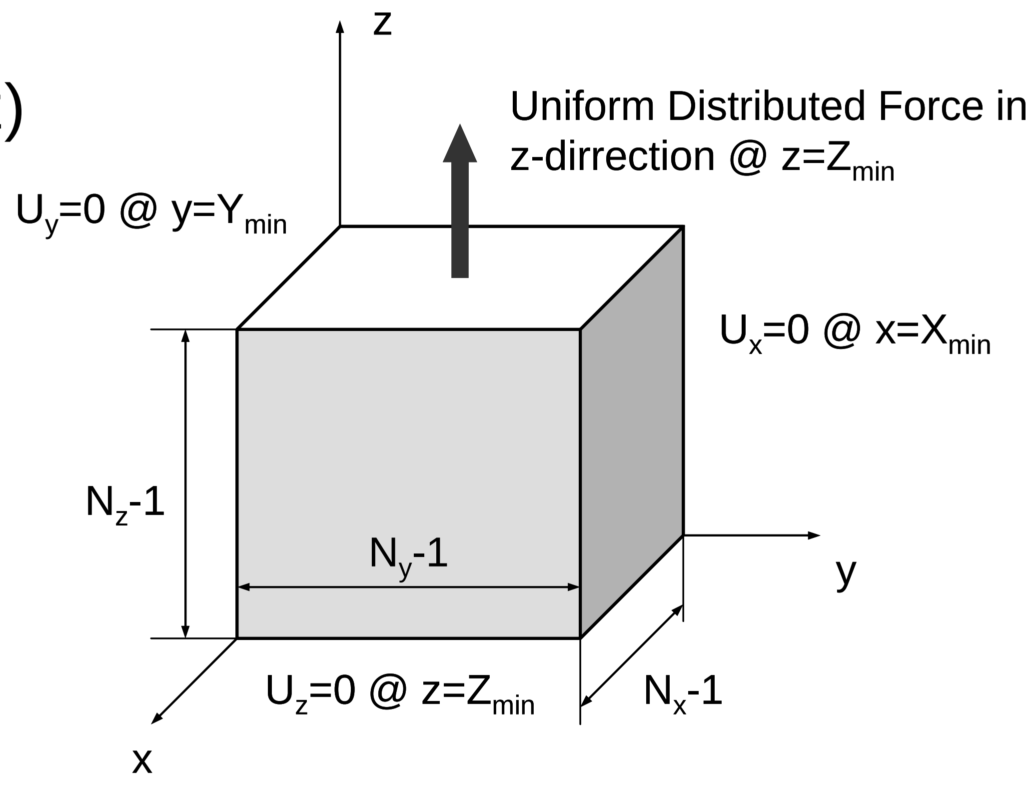




Reduction type loop of DCRS is more suitable for cache-based scalar processor because of its localized operation.

"CUBE" Benchmark

- 3D linear elastic applications on cubes for a wide range of problem size.
- Hardware
  - Single CPU
  - Earth Simulator
  - AMD Opteron (1.8GHz)



Time for $3 \times 64^3 = 786{,}432$ DOF

		DCRS sec. (MFLOPS)	**DJDS original** sec. (MFLOPS)
ES **8.0 GFLOPS**	Matrix	28.6 (291)	**34.2 (240)**
	Solver	360 (171)	**21.7 (3246)**
	Total	389	55.9
Opteron 1.8GHz **3.6 GFLOPS**	Matrix	**10.2 (818)**	12.4 (663)
	Solver	**225 (275)**	271 (260)
	Total	235	283

Time for $3x64^3$=786,432 DOF

		DCRS sec. (MFLOPS)	**DJDS original** sec. (MFLOPS)	**DJDS improved** sec. (MFLOPS)
ES **8.0 GFLOPS**	Matrix	28.6 (291)	**34.2 (240)**	**12.5 (643)**
	Solver	360 (171)	21.7 (3246)	21.7 (3246)
	Total	389	55.9	34.2
Opteron 1.8GHz **3.6 GFLOPS**	Matrix	**10.2 (818)**	12.4 (663)	21.2 (381)
	Solver	225 (275)	271 (260)	271 (260)
	Total	235	283	292

Performance

- Solver
  - Sparse MATVEC
  - DAXPY
  - Dot Products
  - Preconditioning

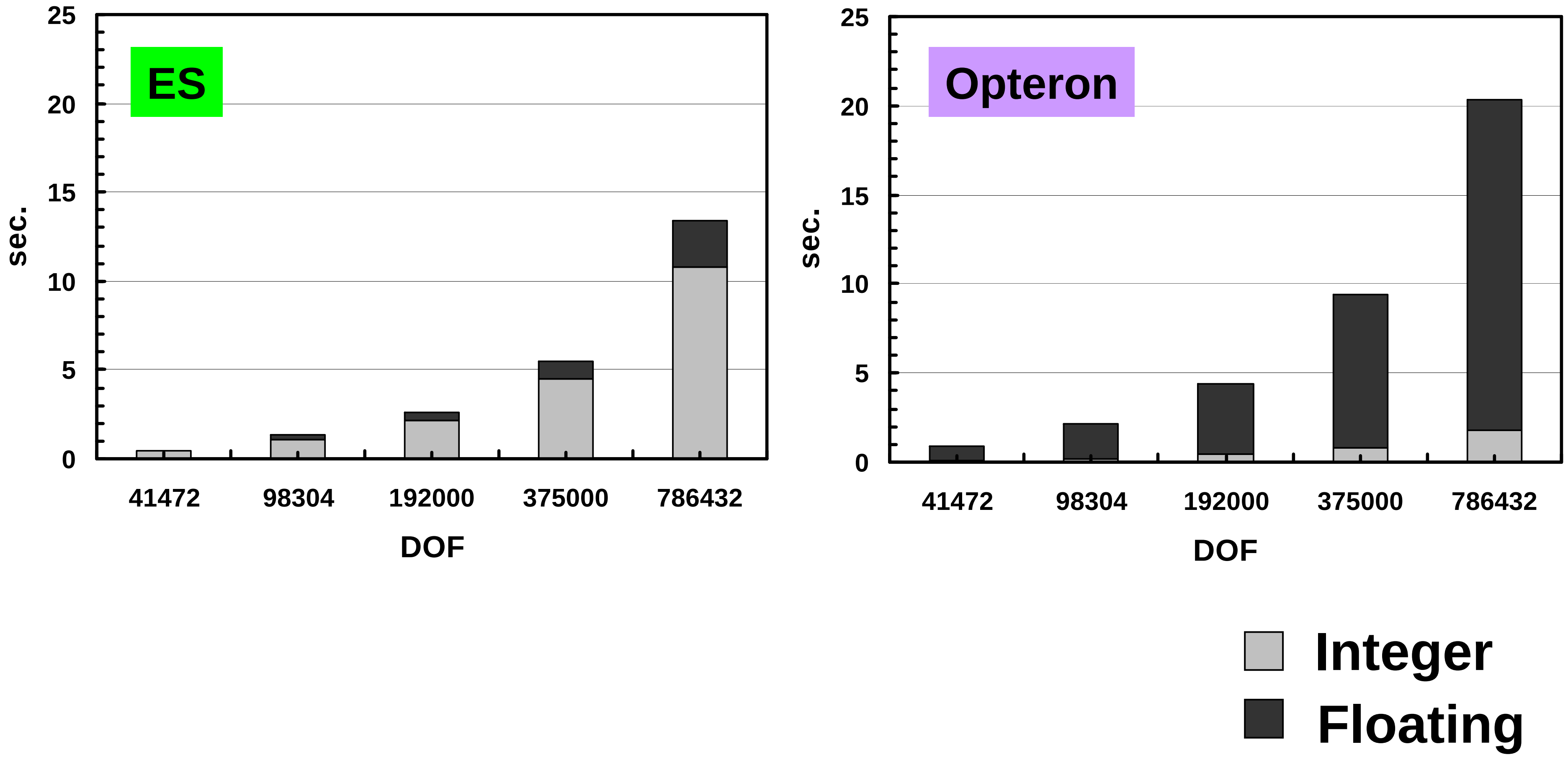
- Matrix
  - Integer Operations
    - connectivity search
    - reordering
  - Floating Operations
    - matrix assembling
    - global accumulation

Performance

- Solver
  - Sparse MATVEC — good for vector processors
  - DAXPY — good for vector processors
  - Dot Products — good for vector processors
  - Preconditioning — good for vector processors

- Matrix
  - Integer Operations — difficult to be vectorized
    - connectivity search
    - reordering
  - Floating Operations — good for vector processors
    - matrix assembling
    - global accumulation

"Matrix" computation time for improved version of DJDS



Suppose “virtual” mode where …

- On scalar processor
  - “Integer” operation part
- On vector processor
  - “floating” operation part
  - linear solvers
- Scalar performance of ES (500MHz) is smaller than that of Pentium III

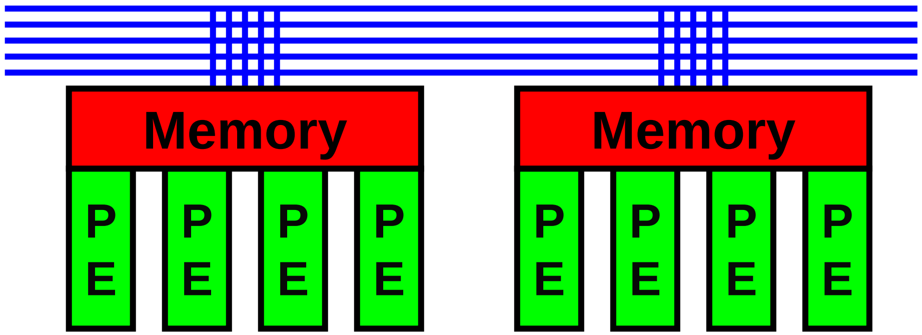
Time for $3x64^3$=786,432 DOF

		DCRS sec. (MFLOPS)	DJDS improved sec. (MFLOPS)	DJDS virtual sec. (MFLOPS)
ES 8.0 GFLOPS	Matrix	28.6 (291)	**12.5 (643)**	**1.88 (4431)**
	Solver	360 (171)	21.7 (3246)	21.7 (3246)
	Total	389	34.2	23.6
Opteron 1.8GHz 3.6 GFLOPS	Matrix	**10.2 (818)**	21.2 (381)	
	Solver	225 (275)	271 (260)	
	Total	235	292	

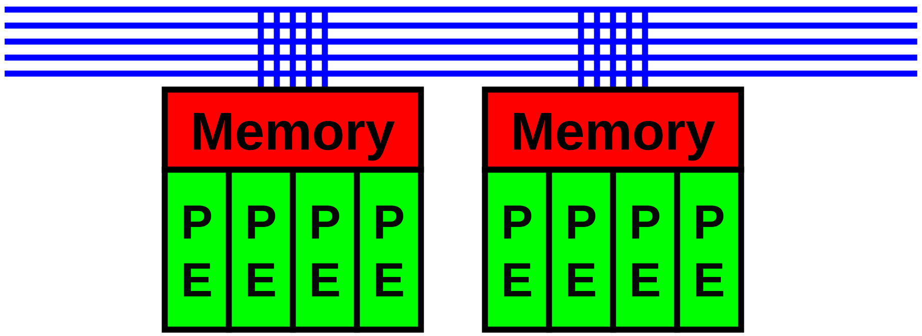
- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- Overview of the Current Project by JST
  - Integrated Predictive Simulation System for Earthquake and Tsunami Disaster
- **Some Technical Issues**
  - Parallel Preconditioning Methods
  - **Vector vs. Scalar Processors**
  - **<u>Parallel Programming Models in Multi-Core Era</u>**
- Future Directions

Flat MPI vs. Hybrid

Flat-MPI：Each PE -> Independent



Hybrid：Hierarchal Structure

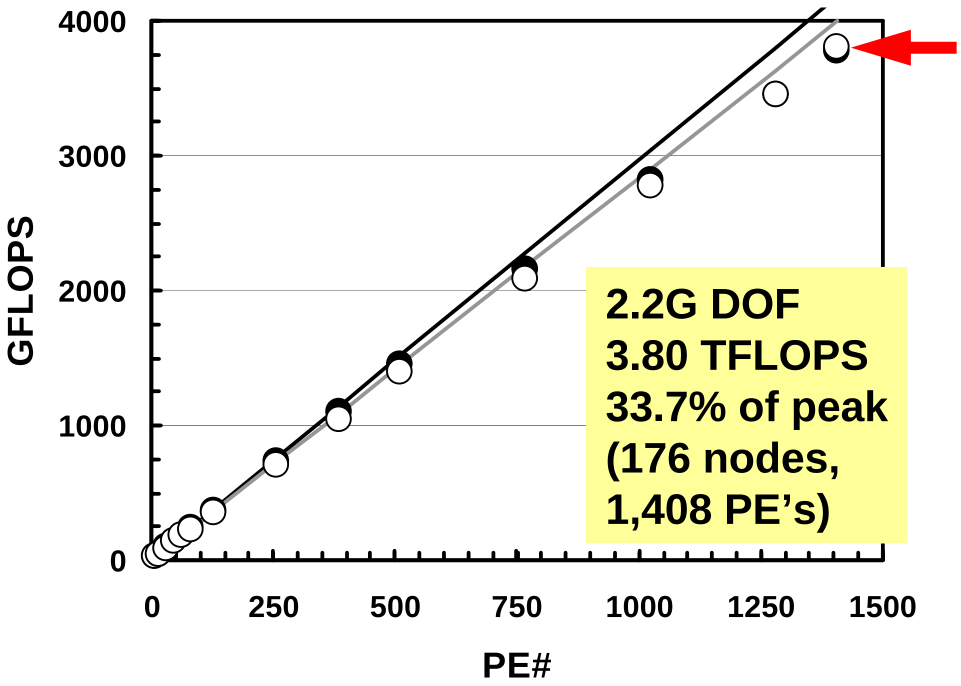


Weak Scaling: LARGE

● Flat-MPI DJDS
○ Hybrid DJDS
— Flat-MPI(ideal)
— Hybrid (ideal)

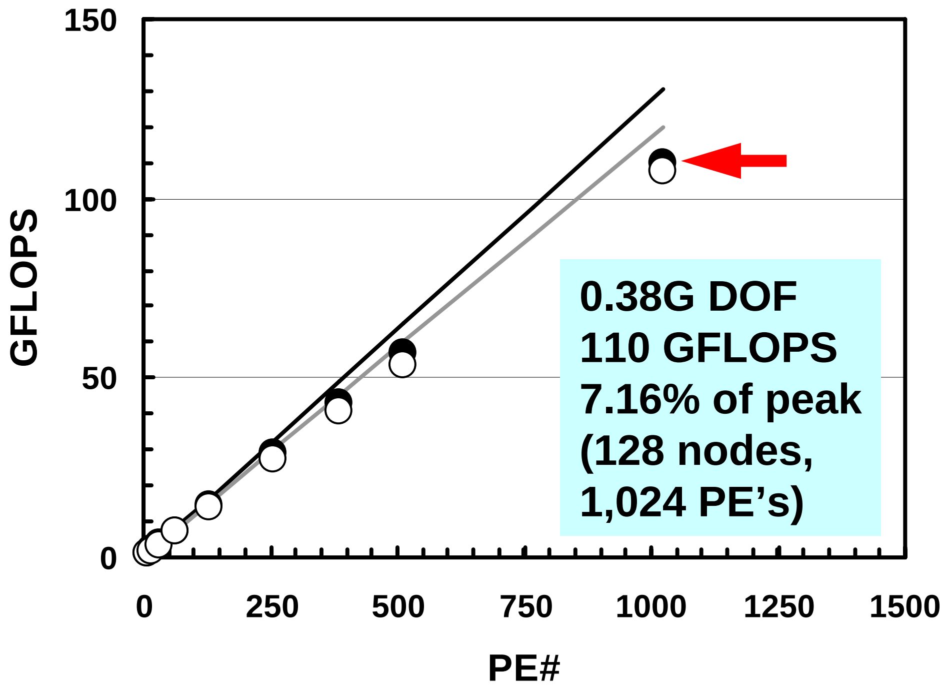
Earth Simulator

1,572,864 DOF/PE
(=3x128x64x64)



IBM SP-3 (Seaborg at LBNL)

375,000 DOF/PE
(=3x50^3)

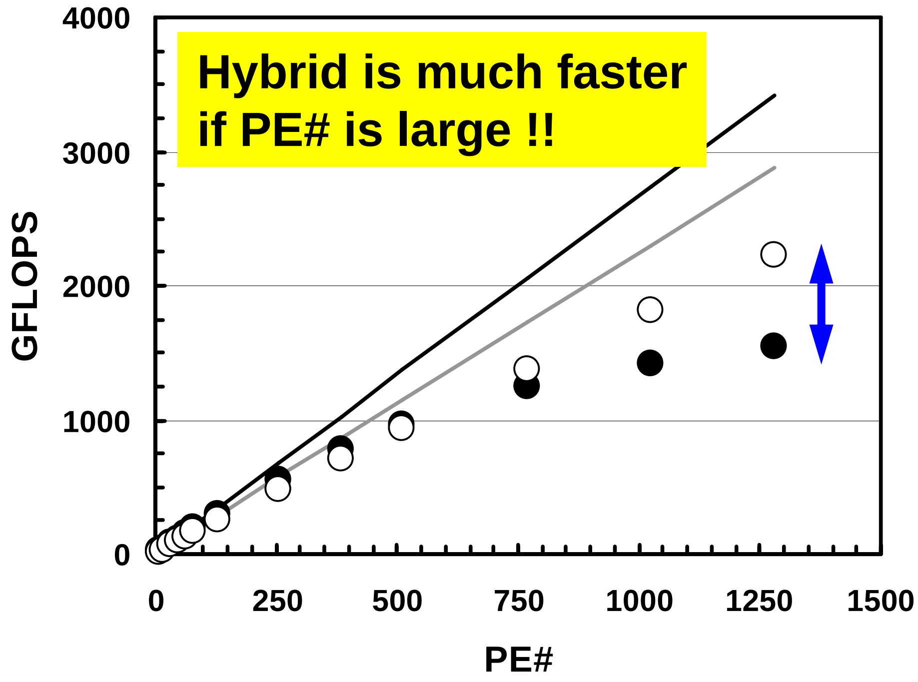


Weak Scaling: SMALL

● Flat-MPI DJDS
○ Hybrid DJDS
— Flat-MPI(ideal)
— Hybrid (ideal)

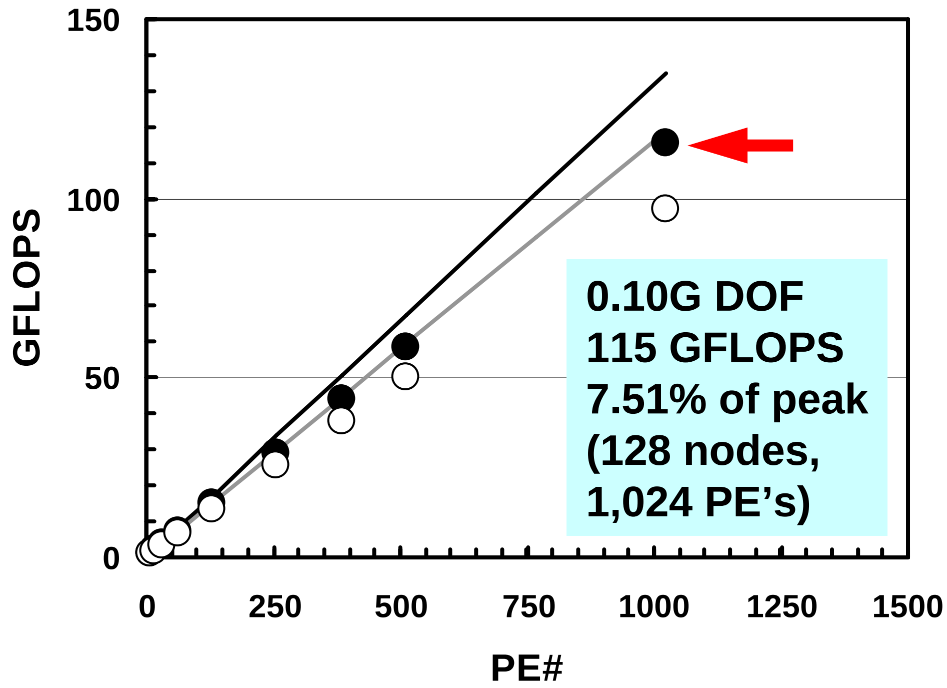
Earth Simulator

98,304 DOF/PE
(=3x32^3)



IBM SP-3 (Seaborg at LBNL)

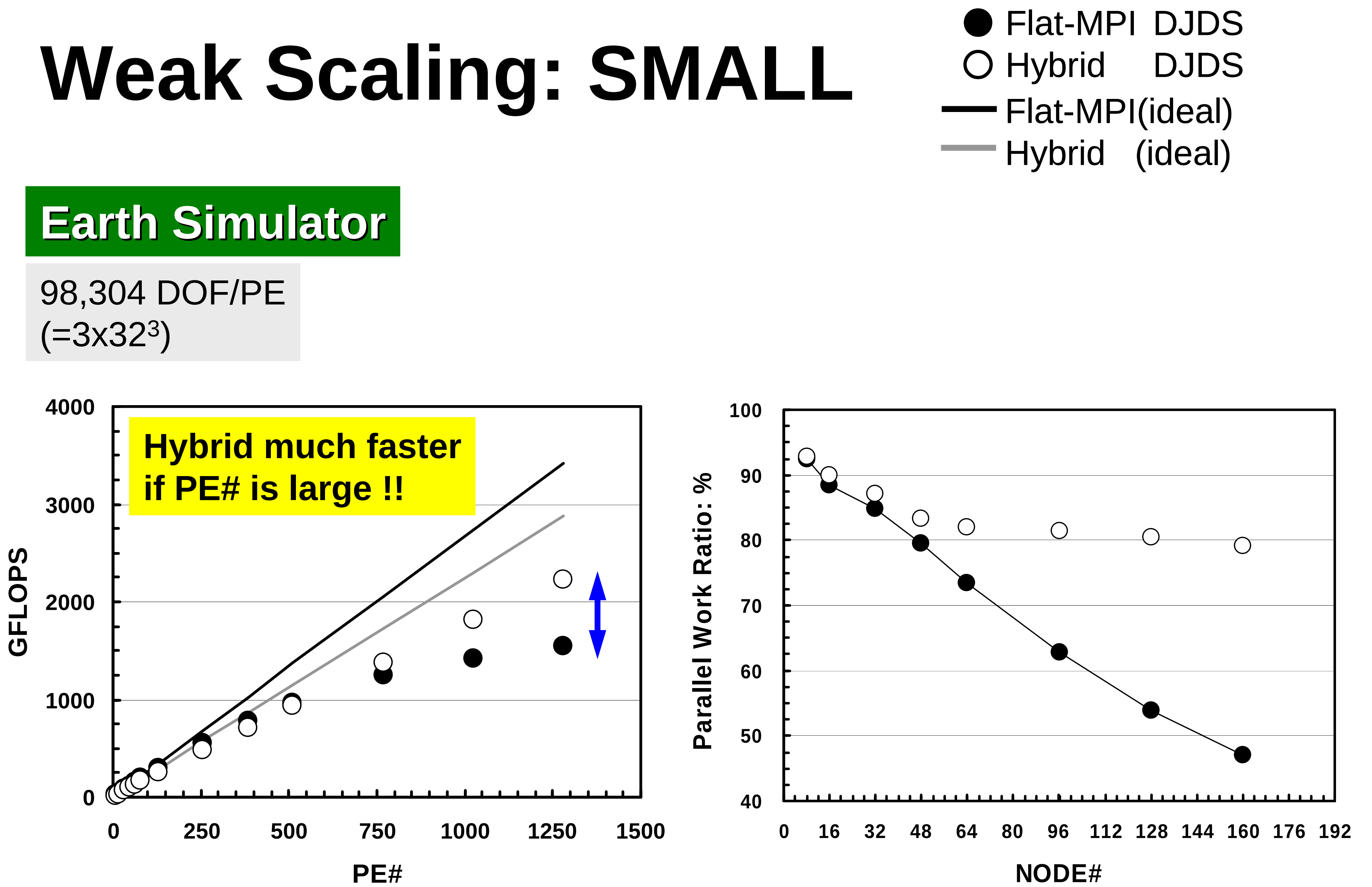
98,304 DOF/PE
(=3x32^3)



Weak Scaling: SMALL

Earth Simulator

98,304 DOF/PE
(=3x32^3)



Platforms

	Earth Simulator	IBM-SP3 (LBNL)
PE#/node	8	16
Clock rate (MHz)	500	375
Peak Performance (GFLOPS/PE)	8.00	1.50
Memory Size (GB/node)	16	16~64
Peak Memory BW (GB/sec/node)	256	16
Network Topology	single stage crossbar	Switch
Network BW (GB/sec/node)	12.3	1.0
MPI Latency (μsec)	5.6-7.7	16.3

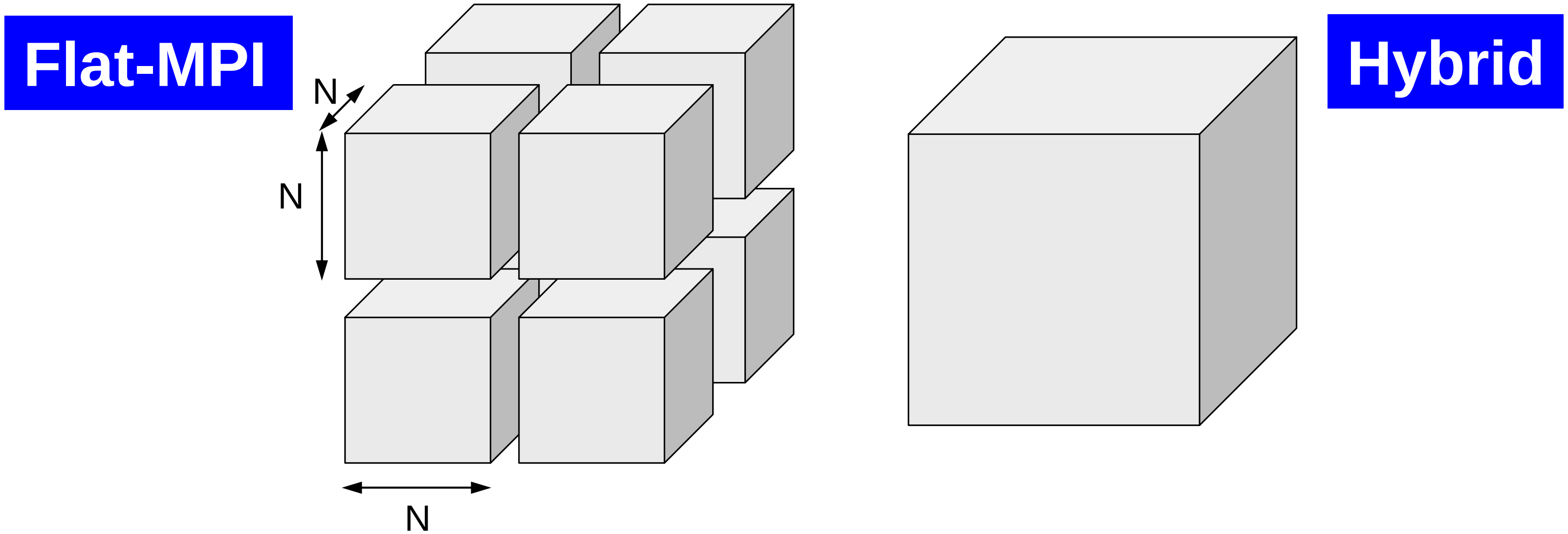
24 : 1*

***based on actual performance**

12 : 1

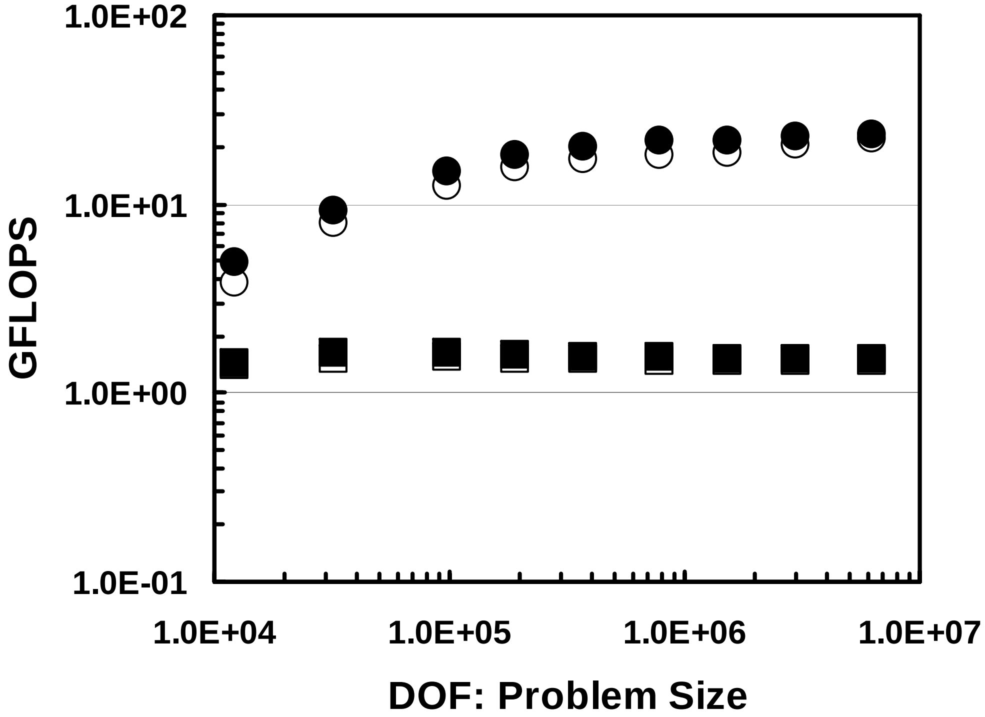
3 : 1

Flat-MPI and Hybrid



	Flat MPI	Hybrid
Problem size/each MPI Process (N=number of FEM nodes in one direction of cube geometry	$3 \times N^3$	$3 \times 8 \times N^3$
Size of messages on each surfaces with each neighboring domain	$3 \times N^2$	$3 \times 4N^2$
Ratio of communication/computation	**1/N**	**1/(2N)**

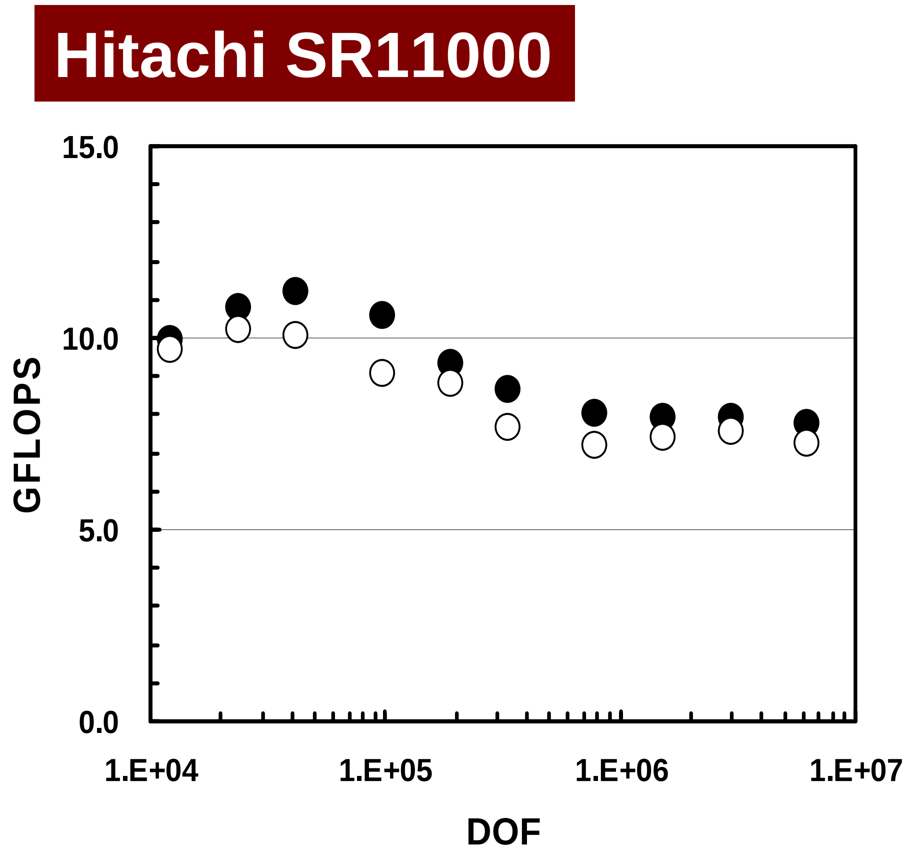
Performance on a Single SMP Node (8 PE's) of the Earth Simulator

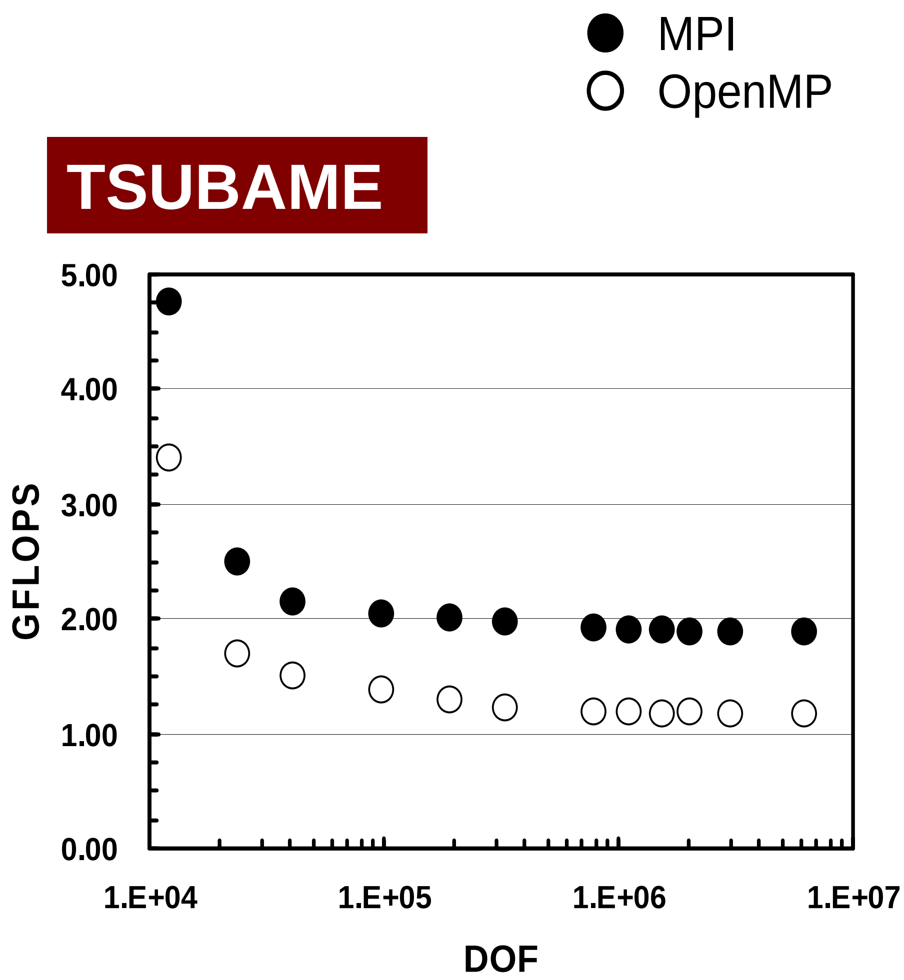


● Flat-MPI DJDS, ○ Hybrid DJDS
■ Flat-MPI DCRS, □ Hybrid DCRS

MPI vs. OpenMP on a Single SMP node

3D Elastic Problems, 8 cores/threads, SGS-CG



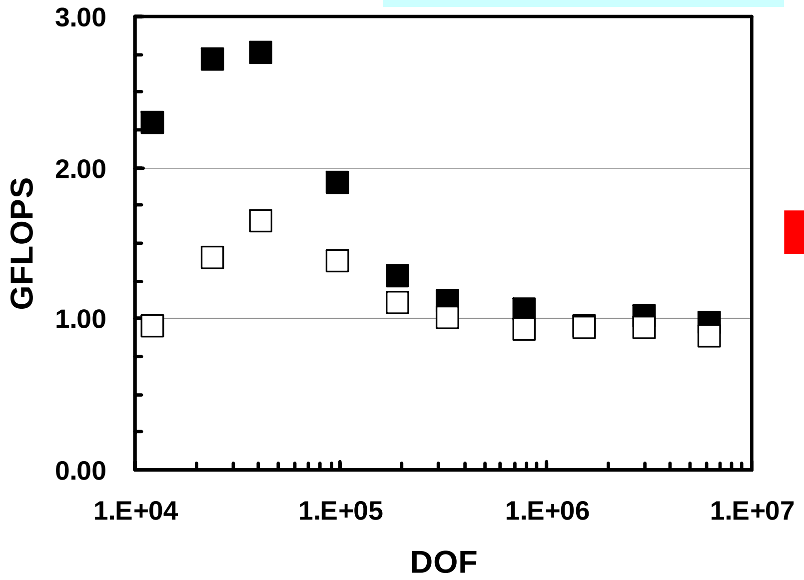


Improvement of Hybrid (OpenMP) Performance on IBM SP-3 ⇒ p5

■ Flat-MPI
□ Hybrid

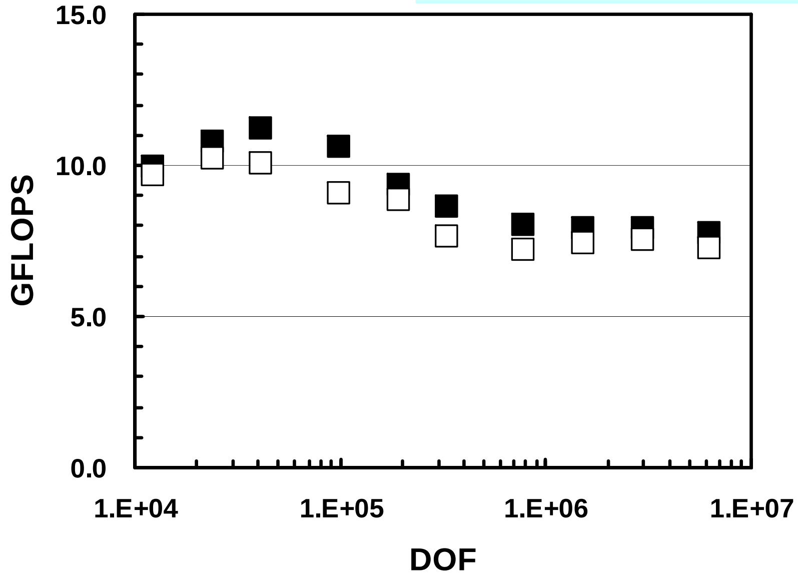
IBM SP-3

375 MHz
1.0 GB/sec
8M L2 cache/PE



IBM p5-595

1.9 GHz
6.4 GB/sec
18M L3 cache/PE



Hybrid vs. Flat MPI

- Generally speaking, “Hybrid” is better for large number of SMP nodes, cores [Nakajima, 2003]

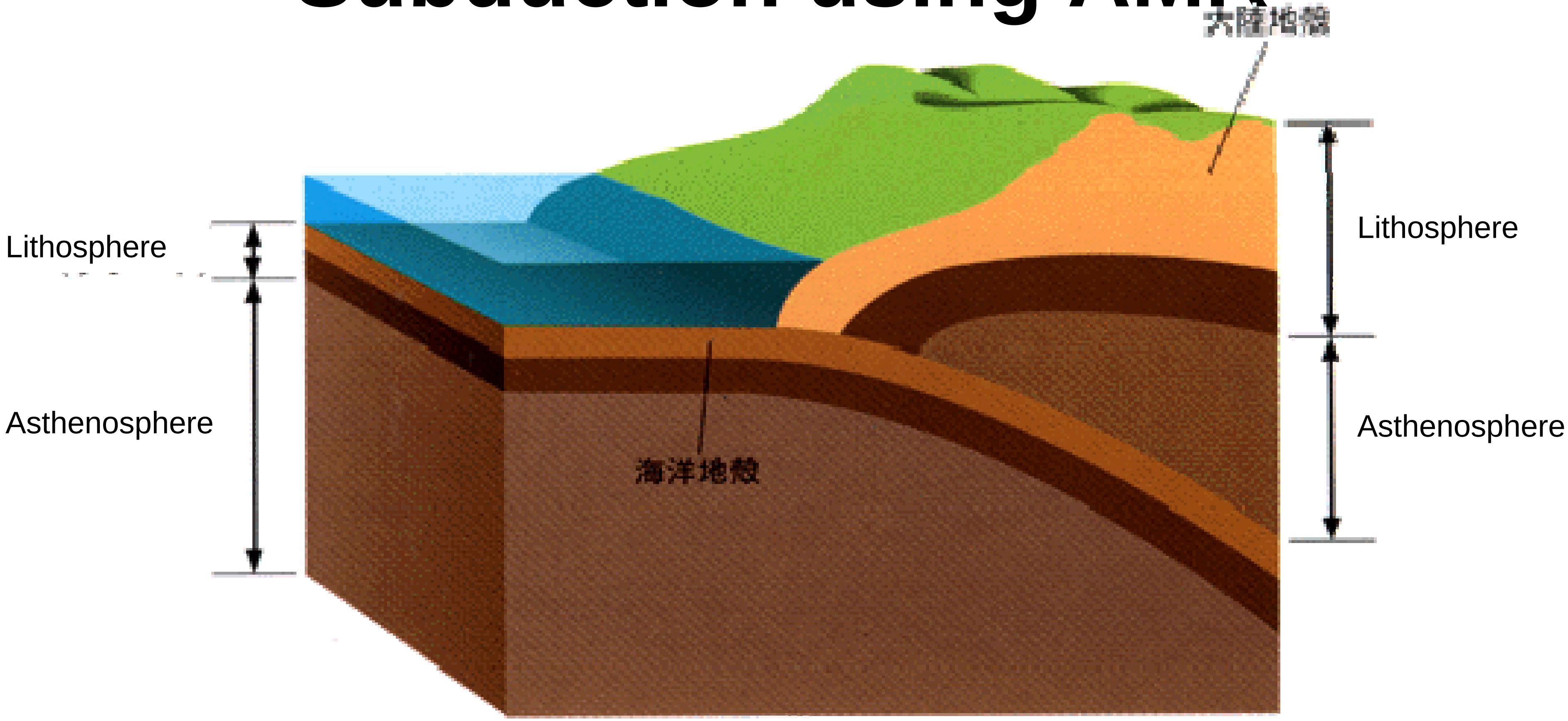
- If single node performance is considered, “Flat MPI” is generally better for most of existing architectures.
  - Performance of Memory
  - The difference is becoming smaller

- Background
  - GeoFEM, HPC-MW
  - COE Program, University of Tokyo
- Overview of the Current Project by JST
  - Integrated Predictive Simulation System for Earthquake and Tsunami Disaster
- Some Technical Issues
  - Parallel Preconditioning Methods
  - Vector vs. Scalar Processors
  - Parallel Programming Models in Multi-Core Era
- Future Directions

Future Directions …

- Many important issues in open discussions
  - Parallel mesh generation
    - AMR, Load Balancing
  - Parallel visualization
    - GeoFEM
  - Data explosion
    - Results of simulations with 10^9 meshes
  - Scalable (O(N)) linear solvers for general ill-conditioned problems
- Future architectures
  - Vector vs. Scalar
    - Special processors for special procedures of numerical algorithms
  - Hybrid vs. Flat MPI
    - It’s too early to make decision.
    - Anyway Flat MPI → Hybrid is not so difficult.

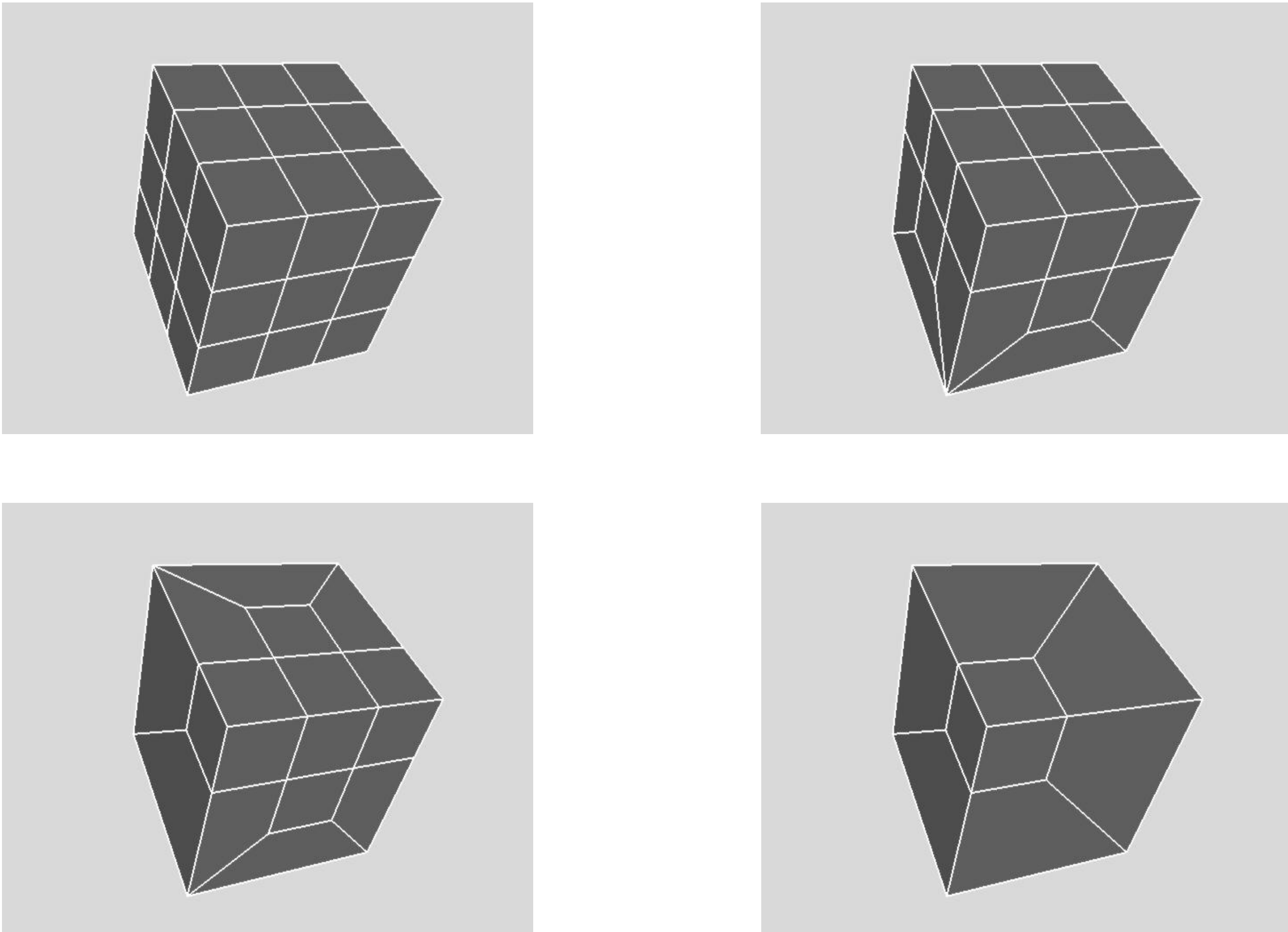
Simulations of Long-Term Plate Subduction using AMR



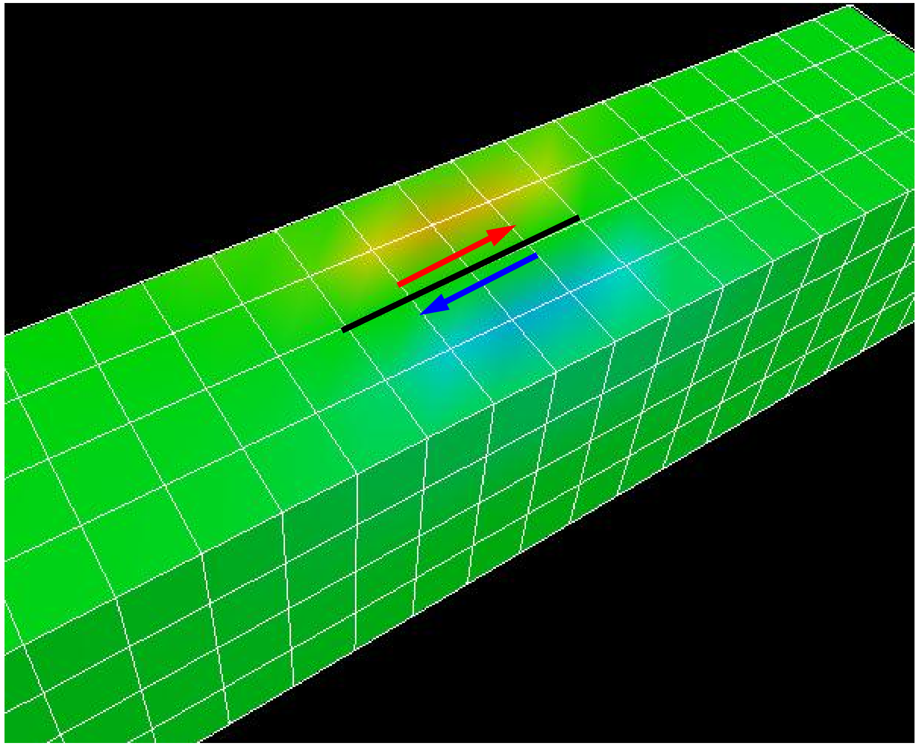
- But people in application side want to use “conforming” octree, because they do not like to change original code/algorithms.

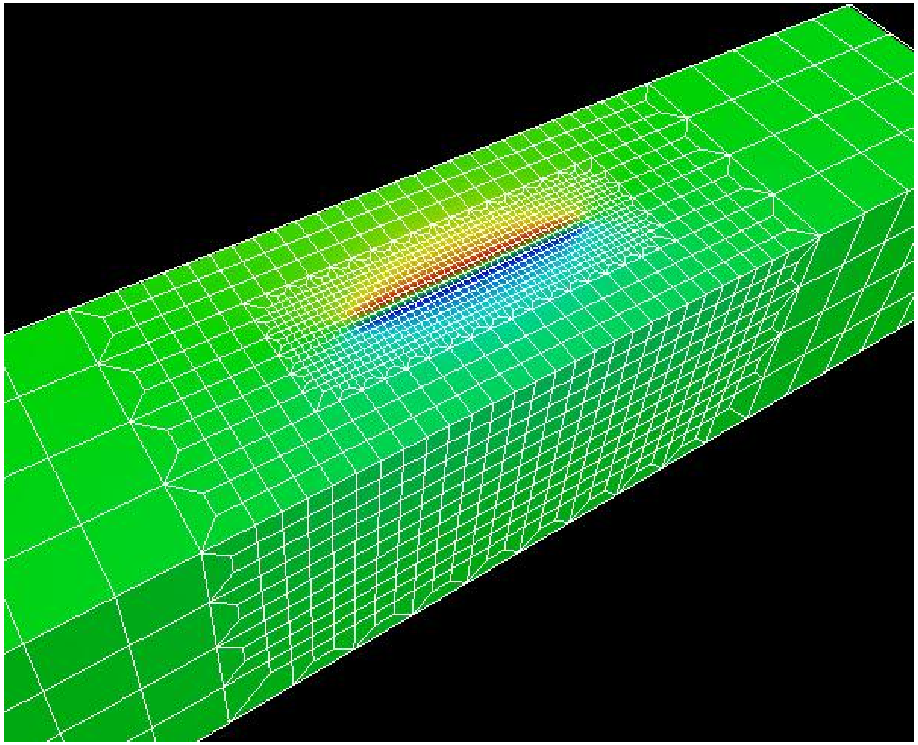
Templates for “conforming” closure of 3D octree structure with adaptive mesh

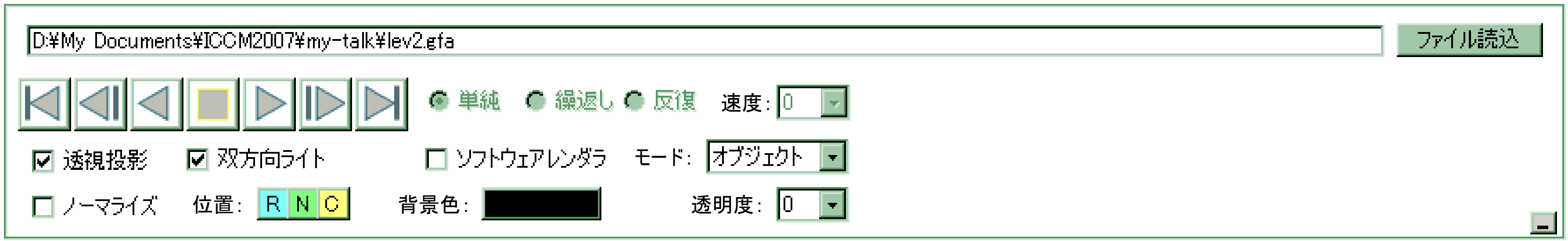
[Schneiders, Shindler & Weiler, 1996]

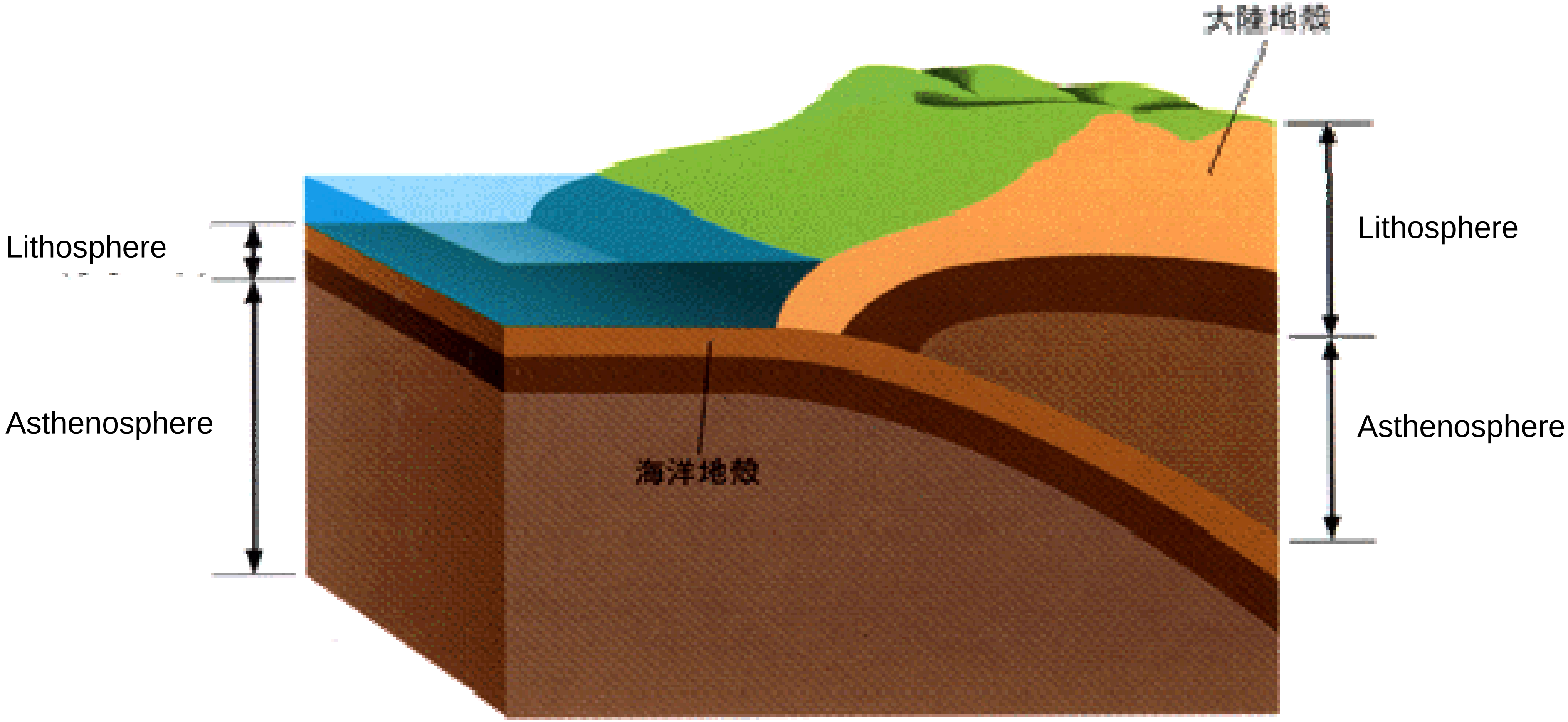


Example






D:¥My Documents¥ICCM2007¥my-talk¥lev2.gfa
ファイル読込
単純
繰返し
反復
速度：
0
透視投影
双方向ライト
ソフトウェアレンダラ
モード：
オブジェクト
ノーマライズ
位置：
R
N
C
背景色：
透明度：
0



- Lithosphere: elastic
- Asthenosphere: visco-elastic
- t → ∞ : Ashenospere→elastic with G→0
  - *Equivalent Theorem* [Fukahata & Matsu'ura, 2006]
  - $\nu \to 0.50$: very ill-conditioned problems: another challenge